

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:36:19 ; Search time 82 Seconds

(without alignments)
118.183 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRONCRAKSSPEEYISTDE.....INNNNNYDNKNNSRDYIK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp Vertebrate:*
- 14: sp Unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2370	99.7	445	10	Q8W5B2 arabidopsis
2	2285	96.1	440	10	Q94CIS arabidopsis
3	923	38.8	631	10	Q93V59 arabidopsis
4	615	25.9	223	10	Q91U50 arabidopsis
5	552	23.2	107	10	Q8W5B1 arabidopsis
6	534.5	22.5	632	10	Q9ZGPO arabidopsis
7	510.5	21.5	692	10	Q9ZNM9 arabidopsis
8	451.5	19.0	186	10	Q94ID6 arabidopsis
9	423	17.8	851	10	Q23524 arabidopsis
10	322.5	13.6	623	10	Q23521 arabidopsis
11	266	11.2	295	10	Q23526 arabidopsis
12	221	9.3	739	4	Q96BD9 homo sapien
13	221	9.3	803	4	Q15022 homo sapien
14	219	9.2	855	5	Q8T9D8 drosophila
15	219	9.2	900	5	Q9NU69 drosophila
16	219	9.2	955	5	Q9W55 drosophila

17	180	7.6	388	10	Q91U51 arabidopsis
18	168	7.1	1231	5	Q97140 dictyostell
19	156.5	6.6	369	5	Q8T1D9 dictyostell
20	153.5	6.5	1309	5	Q8T2H9 dictyostell
21	147.5	6.2	4550	5	Q77336 plasmodium
22	143	6.0	530	5	Q8T2A0 dictyostell
23	141.5	6.0	898	5	Q8T1G7 dictyostell
24	140.5	5.9	548	5	Q8T1S2 dictyostell
25	140.5	5.9	1204	5	Q8T134 dictyostell
26	140.5	5.9	1711	5	Q77322 plasmodium
27	139.5	5.9	1844	5	Q97287 plasmodium
28	138	5.8	682	5	Q9GPR8 dictyostell
29	136	5.7	855	5	Q8SSV6 dictyostell
30	136	5.7	1318	5	Q95PM4 dictyostell
31	136	5.7	1887	5	Q8SSS8 dictyostell
32	135	5.7	164	10	Q9XEF7 Oryza sativ
33	135	5.7	1208	5	Q97101 dictyostell
34	134	5.6	261	5	Q8T868 dictyostell
35	134	5.6	1529	5	Q9GOC2 dictyostell
36	133.5	5.6	1364	5	Q8T223 dictyostell
37	133.5	5.6	3848	5	Q76737 dictyostell
38	133	5.6	1903	5	Q9U0N7 plasmodium
39	131.5	5.5	1245	5	Q96195 plasmodium
40	131.5	5.5	2472	5	Q8T2M5 dictyostell
41	131	5.5	340	5	Q8T212 dictyostell
42	131	5.5	1407	5	Q8T1T5 dictyostell
43	131	5.5	2994	5	Q95ZG5 dictyostell
44	130.5	5.5	699	5	Q77318 plasmodium
45	130.5	5.5	721	5	Q8T218 dictyostell

ALIGNMENTS

RESULT 1	
Q8W5B2	PRELIMINARY; PRT; 445 AA.
AC	Q8W5B2:
DT	01-MAR-2002 (TREMBLrel. 20, Created)
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE	Vernalization 2 protein.
GN	VRN2.
OS	Arabidopsis thaliana (Mouse-ear cress).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC	eucotsid II; Brassicales; Brassicaceae; Arabidopsi.
OX	NCBI_TaxID=3702;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21575675; PubMed=11719192;
RA	Gendall A.R., Levy Y.Y., Willson A., Dean C.;
RT	"The VERNALIZATION2 (VRN2) gene mediates the epigenetic regulation of
RT	vernalization in Arabidopsis.";
RL	Cell 107:525-535(2001).
DR	EMBL; AF284500; AAL32135.1; -
DR	InterPro; IPR000822; Znf_C2H2.
DR	PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
SO	SEQUENCE 445 AA; 51177 MW; 30A3391CE98D877C CRC64;
Query Match	
Best Local Similarity 99.7%; Score 2370; DB 10; Length 445;	
Matches 444; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MCRONCRAKSSPEEYISTDENLLYCKVRLYNIFHLRSLGNPSFLPCLNYKIGAKRR 60
DB	1 MCRONCRAKSSPEEYISTDENLLYCKVRLYNIFHLRSLGNPSFLPCLNYKIGAKRR 60
QY	61 KSRSTGMVYFNKQNNNTLOKTEVREDCSCPCSLGSGFGLOPHLNSHDLDEFEERKL 120
DB	61 KSRSTGMVYFNKQNNNTLOKTEVREDCSCPCSLGSGFGLOPHLNSHDLDEFEERKL 120
QY	121 FEETQTVNYSVLNSFIEEGSDDDKEPEPSLCKPKRRKRGGRNNTRLKVCFLPLD 180

Db 121 SEEQVAVSVKLSNFIIEEGSDDDKFEFPLSCSPRRKRGKGNRRKLVKVFLLD 180
 QY 181 SPSTLNGTENGITLLNGNNGGLGPEATLACGFEEMTSNIPPAIAHSSLDGAKVLTSE 240
 Db 181 SPSTLNGTENGITLLNGNNGGLGPEATLACGFEEMTSNIPPAIAHSSLDGAKVLTSE 240
 QY 241 AVPAATKTRKLSAERSEARSHLLQKOPFYSHRVOPALQVMSDRDSEDDVDYADF 300
 Db 241 AVPAATKTRKLSAERSEARSHLLQKOPFYSHRVOPALQVMSDRDSEDDVDYADF 300
 QY 301 EDROMLDPEVDVKNDEKOPMHLNNSFVRKORVIADGHIISWACEAFSFRYEKEIHRYSLE 360
 Db 301 EDROMLDPEVDVKNDEKOPMHLNNSFVRKORVIADGHIISWACEAFSFRYEKEIHRYSLE 360
 QY 361 WCMRFLIKLNMHGLVDSATINNCNTILENCRNSDDTTNNNSVDPSDSTNNNTIV 420
 Db 361 WCMRFLIKLNMHGLVDSATINNCNTILENCRNSDDTTNNNSVDPSDSTNNNTIV 420
 QY 421 DHPNDINKNVNDKNNNSRDKYIK 445
 Db 421 DHPNDINKNVNDKNNNSRDKYIK 445

RESULT 2
 ID 094CF5 PRELIMINARY; PRT: 440 AA.
 AC 094CF5:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE Hypothetical 50.6 kDa protein.
 GN d14450w.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 RX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,
 RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,
 RA Lam B., Lin J., Meyers R.C., Miranda M., Narusaka M., Nguyen M.,
 RA Palm C.J., Sakurai T., Satou M., Seki M., Shim P., Southwick A.,
 RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.,
 RT Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.
 RU [2]
 RP SEQUENCE FROM N.A.
 RA Tamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Yu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Chek R.,
 RA Kawai J., Kim C., Koesema E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narusaka M., Nguyen M., Palm C., Sakurai T., Satou M., Seki M.,
 RA Shinozaki K., Southwick A., Theologis A., Davis R.W., Ecker J.R.,
 RT Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 RU [2]
 RP SEQUENCE FROM N.A.
 RA EMBL: AY034902; AAK59409.1;
 DR EMBL: AY034902; AAK59409.1;
 DR InterPro: IPR000822; Znf_C2H2.
 DR SMART: SM00355; Znf_C2H2.1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; UNKNOWN.1.
 KW DNA-binding; Hypothetical protein; Zinc-finger.
 SQ SEQUENCE 440 AA: 50623 MW: 0E0AB2C5517BBEF CRC64;

Query Match
 Best Local Similarity 96.1%; Score 2285; DB 10; Length 440;
 Matches 428; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

QY 1 MCRONCRKSSPEEVISTDENLLYCKPVLYNIIFHLRSLGNPSFLPCLNKKIGAKRR 60
 Db 1 MCRONCRKSSPEEVISTDENLLYCKPVLYNIIFHLRSLGNPSFLPCLNKKIGAKRR 60
 QY 61 KSRSTGVAVVNYKDCNNNTLOKTEVRDSCPCSMSCSFGIGLOFHLNNSHDLFEPEFKI 120
 Db 61 KSRSTGVAVVNYKDCNNNTLOKTEVRDSCPCSMSCSFGIGLOFHLNNSHDLFEPEFKI 120
 QY 121 FEEQVAVSVKLSNFIIEEGSDDDKFEFPLSCSPRRKRGKGNRRKLVKVFLLD 180
 Db 121 FEEQVAVSVKLSNFIIEEGSDDDKFEFPLSCSPRRKRGKGNRRKLVKVFLLD 180
 QY 181 SPSTLNGTENGITLLNDGNRGLGPEATLACGFEEMTSNIPPAIAHSSLDGAKVLTSE 240
 Db 181 SPSTLNGTENGITLLNDGNRGLGPEATLACGFEEMTSNIPPAIAHSSLDGAKVLTSE 240
 QY 241 AVPAATKTRKLSAERSEARSHLLQKOPFYSHRVOPALQVMSDRDSEDDVDYADF 300
 Db 241 AVPAATKTRKLSAERSEARSHLLQKOPFYSHRVOPALQVMSDRDSEDDVDYADF 300
 QY 301 EDROMLDPEVDVKNDEKOPMHLNNSFVRKORVIADGHIISWACEAFSFRYEKEIHRYSLE 360
 Db 301 EDROMLDPEVDVKNDEKOPMHLNNSFVRKORVIADGHIISWACEAFSFRYEKEIHRYSLE 360
 QY 361 WCMRFLIKLNMHGLVDSATINNCNTILENCRNSDDTTNNNSVDPSDSTNNNTIV 420
 Db 361 WCMRFLIKLNMHGLVDSATINNCNTILENCRNSDDTTNNNSVDPSDSTNNNTIV 420
 QY 421 DHPNDINKNVNDKNNNSRDK 442
 Db 421 DHPNDINKNVNDKNNNSRDK 440

RESULT 3
 ID 093V59 PRELIMINARY; PRT: 631 AA.
 AC 093V59:
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DE Arabidopsis flower 2.
 GN EMB2.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC euroids II; Brassicales; Brassicaceae; Arabidopsids.
 RX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshida N., Yanai Y.,
 RT Arabidopsis, embryonic flower 2 gene.
 RT Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE FROM N.A.
 RA Yoshida N., Yanai Y.,
 RT Arabidopsis flower 2.
 RT Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB053171; BAB58957.1;
 DR EMBL: AB053171; BAB58956.1;
 DR InterPro: IPR000822; Znf_C2H2.
 DR SMART: SM00355; Znf_C2H2.1.
 DR PROSITE: PS00028; ZINC_FINGER_C2H2.1; UNKNOWN.1.
 KW DNA-binding; Zinc-finger.
 SQ SEQUENCE 631 AA: 71680 MW: F36FE92D0F62E610 CRC64;

Query Match
 Best Local Similarity 38.8%; Score 923; DB 10; Length 631;
 Matches 216; Conservative 52; Mismatches 108; Indels 244; Gaps 9;

QY 1 MCRONCRKSSPEEVISTDENLLYCKPVLYNIIFHLRSLGNPSFLPCLNKKIGAKRR 60
 Db 21 MCRONCRKSSPEEVISTDENLLYCKPVLYNIIFHLRSLGNPSFLPCLNKKIGAKRR 60

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yoshida N., Yanai Y.;
 RT "Arabidopsis, embryonic flower 2 like 2 gene";
 RL Submitted (JAN-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AB053265; BAB58960.1; -
 SO SEQUENCE 186 AA; 21754 MW; D629EB63E9381853 CRC64;

Query Match 19.0%; Score 451.5; DB 10; Length 186;
 Best Local Similarity 50.8%; Pred. No. 1.3e-29;
 Matches 99; Conservative 32; Mismatches 43; Indels 21; Gaps 7;

OY 237 LTSEAVPATKTRKLSAERSEARSHLL-LQKQFYHSHRVQPMALQVMSDSDSEVD- 294
 DB 6 LTTEAKVP-----AKRSKATSHYLP LHKRFYHSRTGQPLSLQVMSDSDSEVDYDK 57
 OY 295 -DDVADFEEDROMLDFVDVKN-DEKQFMHLNMFYRKQVADGHSNACEAFSPRYEKE 352
 DB 58 NDDAAHLESQMLNCSMDENIYAEFFIKLNSFVKQRIYADAHIPNACEAFSLHIOE 117
 OY 333 LHRYSLEFWCWLFLIKLNNHGLVDSATINNCNTLLECRNSDPTTNNNSVDPSDS 412
 DB 118 LRSNLSLDCWRFQKMDYGLDRVTMKNKCNFTIYH--NISTNDINDINNNT-----R 170
 OY 413 MTNNNNIYDHPNDIN 427
 DB 171 TTDNMVDVD--DDIN 183

RESULT 9

023524 PRELIMINARY; PRT; 851 AA.
 AC 023524; 023525;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE Hypothetical 96.8 kDa protein.
 GN ATG16840.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Puidomenech P., Hatzopoulos P., Obermaier B., Duesterhoff A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chaiwatzis N.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 297342; CAB10457.1; -
 DR EMBL: AL161545; CAB80955.1; -
 DR InterPro: IPR002885; PPR: 9.
 DR Pfam: PF01535; PPR: 9.
 DR TIGRFAMs: TIGR00756; PPR: 10.
 KW Hypothetical protein.
 SO SEQUENCE 851 AA; 96849 MW; 9BF5E46E5FEF71C CRC64;

Query Match 17.8%; Score 423; DB 10; Length 851;
 Best Local Similarity 39.6%; Pred. No. 2e-26;
 Matches 101; Conservative 2; Mismatches 0; Indels 152; Gaps 2;

OY 43 PSFLPCLNKTGAKRRK-----SRS 64
 DB 701 PSFLPCLNKTGAKRRKRYAFLLRCSCHSDMLFILLIWKLMNTLSLSRYGHNWMSRS 760

OY 65 TGMVENVKDCNNLTOKTEVRDSCSPCSMLGSPKGLQFHLNSHDLFEFEFLFEY 124
 DB 761 TGMVENVKDCNNLTOKTE----- 779
 OY 125 QTVNVSVKLNSPFIIEEGSDDKFEPFISCKSPRRKRRGGRNNTRLKCPPLDSPSL 184
 DB 780 -----GNRGLEYPEATLQGFEMTSNIPALHSIDAGAKVILTEAVVP 779
 OY 185 TNGTNGITLLDNGRGLGYPATLQGFEMTSNIPALHSIDAGAKVILTEAVVP 244
 DB 780 -----GNRGLEYPEATLQGFEMTSNIPALHSIDAGAKVILTEAVVP 826
 OY 245 ATKTRLSAERSEAR 259
 DB 827 ATKTRLSAERSEAR 841

RESULT 10

023521 PRELIMINARY; PRT; 623 AA.
 AC 023521;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)
 DE Hypothetical 70.1 kDa protein.
 GN ATG16810.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Puidomenech P., Hatzopoulos P., Obermaier B., Duesterhoff A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chaiwatzis N.;
 RL Submitted (JUL-1997) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA EU Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/Genbank/DBJ databases.
 DR EMBL: 297342; CAB10454.1; -
 DR EMBL: AL161545; CAB80952.1; -
 KW Hypothetical protein.
 SO SEQUENCE 623 AA; 70062 MW; 27174171E2C3F03D CRC64;

Query Match 13.6%; Score 322.5; DB 10; Length 623;
 Best Local Similarity 43.4%; Pred. No. 2.9e-18;
 Matches 76; Conservative 28; Mismatches 38; Indels 33; Gaps 6;

OY 237 LTSEAVPATKTRKLSAERSEARSHLL-LQKQFYHSHRVQPMALQVMSDSDSEVD- 294
 DB 6 LTTEAKVP-----AKRSKATSHYLP LHKRFYHSRTGQPLSLQVMSDSDSEVDYDK 57
 OY 295 -DDVADFEEDROMLDFVDVKN-DEKQFMHLNMFYRKQVADGHSNACEAFSPRYEKE 352
 DB 58 NDDAAHLESQMLNCSMDENIYAEFFIKLNSFVKQRIYADAHIPNACEAFSLHIOE 117
 OY 333 LHRYSLEFWCWLFLIKLNNHGLVDSATINNCNTLLECRNSDPTTNNNSVD 407
 DB 118 LRSNLSL-----DLGNTEETETAPV-----ATEATTHGEDGID 151

RESULT 11

023526 PRELIMINARY; PRT; 295 AA.
 AC 023526;
 DT 01-JAN-1998 (TREMblrel. 05, Created)
 DT 01-JAN-1998 (TREMblrel. 05, last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, last annotation update)

Hypothetical 33.6 kDa protein.
 GN ATG16850.
 OS Arabidopsis thaliana (mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OC NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Beyer M., Stiekema W., Murphy G., Wandutt R., Pohl T., Terry N.,
 RA Kreis M., Kavanagh T., Entian K.D., Rieger M., James R.,
 RA Pudgomech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
 RA Jones J., Palme K., Ansgore W., Delseny M., Bancroft I., Mewes H.W.,
 RA Scheller C., Chalevalais N.,
 RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z97342; CAB10459.1;
 DR EMBL; AL161545; CAB80956.1;
 KW Hypothetical protein.
 SQ SEQUENCE 295 AA; 33564 MW; F22BD5E82031423 CRC64;

Query Match 11.2%; Score 266; DB 10; Length 295;
 Best Local Similarity 53.4%; Pred. No. 5, 3e-14;
 Matches 70; Conservative 3; Mismatches 30; Indels 28; Gaps 4;

QY 278 MALEOVMSDSEDDVDVADVEDQ-----MDDFYDVNKKDEQ 318
 DB 1 MALEOVMSDSEDDVDVADVEDQVHDFPLSLSQQRMVYDMDDVDVNNKDEQ 60
 QY 319 FMHLSNFKRQVADGHSWACEAFSPFEKEKRRYSSEFCWRLFLKLMNGLVDS 378
 DB 61 FMHLSNFKRQVADGHSWACEAFSPFEKEKRRYSSEFCWRLFLKLMNGLVDS 111
 QY 379 ATINNCNTILE 389
 DB 112 ALANVHTRE 122

RESULT 12
 Q96BD9 PRELIMINARY; PRT; 739 AA.
 AC Q96BD9;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Joined to JAZF1
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UTERUS;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC015704; AAH15704.1;
 DR InterPro; IPR000822; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2.1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 KW DNA-binding; Zinc-finger.
 SQ SEQUENCE 739 AA; 83054 MW; A8830EBC3FD38D56 CRC64;

Query Match 9.3%; Score 221; DB 4; Length 739;
 Best Local Similarity 21.1%; Pred. No. 9e-10;
 Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

QY 58 RRRK-----SRSTGMVFNKYKDCNNTLQTEVERDCSCPFCSMLGSGFKGLQFHLNNSHD 112
 DB 413 RREKDPENRQKRLRIFYOFLYNNNTROOTEARDDLHCPMCTLNCRKLYSLTKHLKLCHS 472

QY 113 LFEFEKLFEEYQTVNVAVKLNSFFEESSDDDKFEPFSLCSNPK-RROR---GRNN 168
 DB 473 RFFNVPVHPKARIDVSI-----NECIDG-----SYAGNPDIHQPGFARSNG 518
 QY 169 -TRRLKVCPLPDSLSLNGTENGITLLDNGNRGLGPEATELAGOFEWTSNIPALIAS 227
 DB 519 PVKRPFIHTLVCRRKFKASSEFLESDD----- 549
 QY 228 SLDAKAVILISEAVVPATKTRKLSAERSEARSHLLQKQFYHSHRQPMALQVMSDR 287
 DB 550 -----EVEQORTYSSG-----NLTYPHSDCULPRLPOEM--EV 581
 QY 288 DSEDDVDVADVEDQRMIDDFVYNKDEQFMHLSNFKRQVADGHSWACEAFSR 347
 DB 582 DSEDDVDVADVEDQRMIDDFVYNKDEQFMHLSNFKRQVADGHSWACEAFSR 347
 QY 348 FFEKELHRSYSLFCWRLFLKLMNGLVDSATINNCNTILENCRSSPTTTNNNSVD 407
 DB 642 NFGKRIK-KNLGRFMHLVSMHDFNLISMSIDKAVTKLRMQ-----QKLE 689
 QY 408 RPSDSTNNNNIVDHPN-----DINKNNVNDKONS 439
 DB 690 KGSASPAHEETEEQNGTANGFSEINSKEALETDSVS 728

RESULT 13
 Q15022 PRELIMINARY; PRT; 803 AA.
 ID Q15022;
 AC Q15022;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE KIAA0160 protein (Fragment).
 GN KIAA0160.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96127530; PubMed=8590280;
 RA Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;
 RT Prediction of the coding sequences of unidentified human genes. IV.
 RT The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
 RT analysis of cDNA clones from human cell line K5-1.
 RL DNA Res. 2:167-174(1995).
 DR EMBL; D63881; BA009931.1;
 DR InterPro; IPR000822; Znf_C2H2.
 DR SMART; SM00355; Znf_C2H2.1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
 FT NON_TER
 SQ SEQUENCE 803 AA; 89963 MW; CDFB901A35F29A7C CRC64;

Query Match 9.3%; Score 221; DB 4; Length 803;
 Best Local Similarity 21.1%; Pred. No. 1e-09;
 Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

QY 58 RRRK-----SRSTGMVFNKYKDCNNTLQTEVERDCSCPFCSMLGSGFKGLQFHLNNSHD 112
 DB 477 RREKDPENRQKRLRIFYOFLYNNNTROOTEARDDLHCPMCTLNCRKLYSLTKHLKLCHS 536
 QY 113 LFEFEKLFEEYQTVNVAVKLNSFFEESSDDDKFEPFSLCSNPK-RROR---GRNN 168
 DB 537 RFFNVPVHPKARIDVSI-----NECIDG-----SYAGNPDIHQPGFARSNG 582
 QY 169 -TRRLKVCPLPDSLSLNGTENGITLLDNGNRGLGPEATELAGOFEWTSNIPALIAS 227
 DB 583 PVKRPFIHTLVCRRKFKASSEFLESDD----- 613
 QY 228 SLDAKAVILISEAVVPATKTRKLSAERSEARSHLLQKQFYHSHRQPMALQVMSDR 287
 DB 614 -----EVEQORTYSSG-----NLTYPHSDCULPRLPOEM--EV 645

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:17:53 ; Search time 1928 Seconds

(without alignments)
3738.066 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRACKSPEVISTDE.....INNKNVNDKNNRSRDYIK 445

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+g2n_model -DEV=x1h
-O=/cgn2.1/USPPO_spool/US09890220/runat_13062003_144851_16490/app-query.fasta.1.583
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -IOFCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MTELEN=0 -MAXLEN=200000000
-USER=US09890220_ecgn2.1_13062003_144851_16490 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEDUERY -NEG_SCORES=0 -WAIT -DSEBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estp1:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_lin:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	778	32.7	672	10	AV822548
2	760.5	32.0	1128	11	AV104964
3	568	23.9	680	10	AM038171
4	553	23.3	769	14	BO505017
5	545	22.9	846	12	BG648271
6	528	22.2	480	9	AJ469021
7	510.5	21.5	561	14	BQ740672
8	509	21.4	600	13	BI479743
9	506	21.3	391	10	AV817525
10	498.5	21.0	569	14	BO611550
11	456.5	19.2	474	10	AM234600
12	450	18.9	467	12	BG580495
13	448	18.8	405	12	BG314044
14	444.5	18.7	662	14	BQ279630
15	442.5	18.6	612	10	BE203328
16	437	18.4	444	12	BG551103
17	436	18.3	554	13	BI784785
18	422	17.7	429	14	BQ459233
19	420	17.7	546	13	BI321750
20	396.5	16.7	576	10	AM076157
21	394	16.6	364	9	AI164598
22	386	16.2	732	14	BQ841916
23	375	15.8	561	14	BQ163202
24	374	15.7	644	13	BQ453006
25	372	15.6	757	13	BQ460553
26	367.5	15.5	445	10	AM202144
27	362.5	15.2	594	10	AM000259
28	347	14.6	519	14	BQ168835
29	343.5	14.4	326	9	AI163743
30	329	13.8	727	13	BI932726
31	306	12.9	727	14	BQ801572
32	304.5	12.8	560	9	AI179561
33	304	12.8	526	17	CNS00MYB
34	303	12.7	586	14	BQ243123
35	302.5	12.7	529	12	BG459005
36	299.5	12.6	515	12	BR481866
37	274	11.5	677	14	BQ639737
38	271	11.4	322	10	AM309408
39	269	11.3	236	14	BQ904049
40	268	11.3	804	17	BH587461
41	267	11.2	553	12	BG909756
42	267	11.2	553	14	BQ620168
43	257	10.8	523	13	BI992698
44	256.5	10.8	632	12	BE821501
45	254.5	10.7	416	10	AM432863

ALIGNMENTS

RESULT 1
AV822548
LOCUS
DEFINITION AV822548 RAP15 Arabidopsis thaliana CDNA clone RAFL05-08-K19 5',
ACCESSION
VERSION AV822548.1 GI:19864597
KEYWORDS
SOURCE
ORGANISM
thale cress.
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1 (bases 1 to 672)
Seki, M., Narusaka, M., Ishida, D., Kamuya, A., Satou, M., Nakajima, M.,

Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arkawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)
 Contact: Motoko Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msekier@r.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). This clone is in a
 modified pBluescript vector as a SstI/XhoI insert. Please visit our
 web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for
 further details.

FEATURES

source

1. 672
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /clone="RAF105-08-R19"
 /clone_1db="RAF15"
 /dev_stage="rossette plants"
 /lab_host="SOLR"
 /note="Site_1: SstI, Site_2: XhoI; subjected to
 dehydration-treated(1,2,3,10,24 hr)"
 BASE COUNT 197 a 116 c 143 g 216 t
 ORIGIN

Alignment Scores:

Pred. NO.: 6.11e-80 Length: 672
 Score: 778.00 Matches: 144
 Percent Similarity: 99.32% Conservative: 1
 Best Local Similarity: 98.63% Mismatches: 0
 Query Match: 32.72% Indels: 0
 DB: 10 Gaps: 0

US-09-890-220-2 (1-445) x AV822548 (1-672)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 |||||
 DB 233 ATGTGTAGGCGAGAAATGTCTCGCGAATCTCTACCGGAGAGAGTGAATTTCACTGATGAG 292
 |||||
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 |||||
 DB 293 AATCTCTGATATATGTAACTGTTCGACTATATAACATCTTCACCTTGCTGCTCTGA 352
 |||||
 QY 41 GlyAspProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 |||||
 DB 353 GGCACCCATCGTTTCTGCCACAGATCTTGAACTACAAATTTGGGCAAGGCAAAAGA 412
 |||||
 QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 |||||
 DB 413 AAGTCAGATCTACTGGAGATGAGTTCACATAAAGATTGTAATTAATCAATTACAA 472
 |||||
 QY 81 LysThrGluValArgLysLysCysSerCysProPheCysSerMetLeuGlySerPhe 100
 |||||
 DB 473 AGAAGCTGAAGTTGGAGAGATTTGTTGTCATTTTGCCTATGCTATGCTAGCTTC 532
 |||||
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 |||||
 DB 533 AAGGGCTGCAATTTCAATTTGAAATTCATCATGATTTTGAATTTGAGTTCAAGCTT 592
 |||||
 QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 |||||
 DB 593 TTGGAGAAATACACAGACATTAATGTTCTGTAAACCTTAATTCCTCATATTTGAGGA 652
 |||||
 QY 141 GluGlySerAspAsp 146
 |||||
 DB 653 GAAAGAGTGAATGATGAT 670

RESULT 2

AY104964
 LOCUS AY104964 1128 bp mRNA linear HTC 25-MAY-2002
 DEFINITION Zea mays PCO105783 mRNA sequence.
 ACCESSION AY104964
 VERSION AY104964.1 GI:21208042
 KEYWORDS
 SOURCE
 ORGANISM
 Zea mays.
 Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PAC
 clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 1128)
 Hainey, C.F., Dolan, M., Miao, G.H., Vogel, J.M., Whitesitt, M.S.,
 Arthur, L.W., Hanafey, M., Morgante, M., and Tingey, S.V.
 Maize Mapping Project/Dupont Consensus Sequences for Design of
 Overgo Probes
 Unpublished (2002)
 2 (bases 1 to 1128)
 Direct Submission
 Coe, E.C.
 Submitted (25-APR-2002) Maize Mapping Project, University of
 Missouri, Columbia, MO 65211, USA
 Location/Qualifiers

FEATURES

source

1. 1128
 /organism="Zea mays"
 /db_xref="taxon:4577"
 /db_xref="taxon:4577"
 /clone="PCO105783"
 /clone_1db="Maize Mapping Project/Dupont Consensus
 Library"
 /note="this sequence is part of a project of EST
 assemblies resulting from the application of public
 configs to seed Dupont contigs; this resource was
 assembled by Dupont as part of a collaboration for the
 overgo addressing of BACs in conjunction with the Maize
 Mapping Project"
 BASE COUNT 355 a 199 c 242 g 332 t
 ORIGIN

Alignment Scores:

Pred. NO.: 1.48e-77 Length: 1128
 Score: 760.50 Matches: 157
 Percent Similarity: 64.35% Conservative: 47
 Best Local Similarity: 49.53% Mismatches: 80
 Query Match: 31.98% Indels: 33
 DB: 11 Gaps: 7

US-09-890-220-2 (1-445) x AY104964 (1-1128)

QY 86 GluAspCysSerCysProPheCysSerMetLeuGlySerPheGlyLeuGlnPhe 105
 |||||
 DB 12 GAAGATTTCTCTGTCATTTGCTATGACATTTGGAAAGCTTCAGAGGTCAGATGC 71
 |||||
 QY 106 HisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluGluTyrGln 125
 |||||
 DB 72 CATTTAACTCATCATGATCATTCACATATGAGTTTGGATATGTAAGAGTACAG 131
 |||||
 QY 126 ThrValAsnValSerValLysLeuAsnSerPhe-----IlePheGluGluGluGly 142
 |||||
 DB 132 GTTGTAAAGTTAGTGTGAAGGCTGATGCTTGAGAACAGACGCTTTTGGCGAG- 185
 |||||
 QY 143 SerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArgGln 162
 |||||
 DB 186 GCGCTGATCCAGAGCATCAACATTTTATGCTCAAGGTTTAAAGACGTAAGACGA 245
 |||||
 QY 163 ---ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSer 181
 |||||
 DB 246 TCAAAAGACCAATGAGGAGAAATCAGGATGACCTCACAATATTATGAAATCAGTTCA 305
 |||||
 QY 182 ProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGly 201
 |||||
 DB 306 CCTGAAGAGAGGAGGATGATGAGACCAACTTTGT-----CAAGGGAGAAATGGG 356

QY 202 LeuGIYTYrProGLuAlaThrGluleuAlaGlyInPheGluMetThrSerAsnIlePro 221
 DB 356 ----- 356
 QY 222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaVal-----IleLeuThr 238
 DB 357 ACTTCTGAGCAAAATGCTTCATTCCTGCTCAATCTTACATGAGCAGCAATCTTCA 416
 QY 239 SerGluAlaValAlaProAlaThrIleThrIleGlyLeuSerAlaGluArgSerGluAla 258
 DB 417 CCACCAACAGTACTACAGTTGGGAGCAAGAGAGCTATCT--GAGAGATCTGACCT 473
 QY 259 ArgSerHisLeuLeuGluInLysArgGlnPheThrIleSerHisArgValGlnPromet 278
 DB 474 AGAATTCGGCAACTCTCCAAAAGACAGATTCCTCCATCTCCACAGGGCCGACCAATG 533
 QY 279 AlaLeuGluInValMetSerAspArgSerGluAspGluValAlaAspAspValAla 298
 DB 534 CAACGTGAGCAAGTGTCTCGGACCGTGTAGTGAAGATGAAGTGTATGATATCTCT 593
 QY 299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAspGluLysGln 318
 DB 594 GACTTCGAGATAGAGATGCTGTATGATTTGTTATGTTACGAAAGATGAAACTT 653
 QY 319 PheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIle 338
 DB 654 ATTAGCATATGTGGAATTCATTTCTTGAAACAAAGCTTGATGCTGATGCTATATA 713
 QY 339 SerTrpAlaGlyGluAlaPheSerArgPheThrGlyLysGluLeuHisArgTrpSer 358
 DB 714 CTTGGGCGCTCGAGGCGATCTCCAGATGATGACGACCAACTATATCAAAATCTCTCT 773
 QY 359 LeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSer 378
 DB 774 CTGCTGTGGGGTGGCTTCTTCATGATTAACCTTGGAACCATACATTAATTAATGCTC 833
 QY 379 AlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
 DB 834 CGCACTATGACACATGCAATACAGTCTCTCAATTTATACAAAGAAAGAAC 884
 RESULT 3
 AM038171 680 bp mRNA linear EST 18-MAY-2001
 LOCUS EST279828 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA
 DEFINITION clone cLET1N5, mRNA sequence.
 ACCESSION AM038171
 VERSION AM038171.1 GI:5896925
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 680)
 D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Jiang, F., Upton, J.,
 Romling, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nieman, W.,
 Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni,
 J.,
 Generation of ESTs from tomato leaf tissue
 Unpublished (1999)
 CONTACT CUGI
 ORGANISM Lycopersicon esculentum
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 5 prime sequence.
 LOCATION/Qualifiers
 1..680
 /organism="Lycopersicon esculentum"
 /cultivar="Rio Grande Ptor"
 /db_xref="taxon:4081"
 /clone="cLET1N5"

/clone.lib="tomato mixed elicitor, BTI"
 /tissue.type="leaf"
 /dev.stage="4-6 week old plants"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
 XhoI; cLET - Inoculated with a variety of disease response
 elicitors. Plants exposed to 2,6 dichloroisonicotinic
 acid, BTH, jasmonic acid, ethylene, fenchone, Etx,
 Okadaic acid, or systemin prior to tissue harvest. EcoRI
 site was destroyed during cloning."
 BASE COUNT 192 a 120 c 164 g 204 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,95e-55 Length: 680
 Score: 568.00 Matches: 107
 Percent Similarity: 84.35% Conservative: 17
 Best Local Similarity: 72.79% Mismatches: 21
 Query Match: 23.89% Indels: 2
 DB: 10 Gaps: 2
 US-09-890-220-2 (1-445) x AM038171 (1-680)
 QY 254 GluArgSerGluAlaArgSerHisLeuLeuGluInLysArgGlnPheThrHisSerHis 273
 DB 4 GAGCGCTGTGATCCAGAAATGCTGACCTCTCGCAAAAAGCAATCTTCTCATCTCAT 63
 QY 274 ArgValGlnPrometAlaLeuGluInValMetSerAspArgAspSerGluAspGluVal 293
 DB 64 AGGGCCGAGCCCATGCTGAGCAAGTGTATGATGACCGAGACAGTGAAGTAACT 123
 QY 294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
 DB 124 CATGATGATGTTGCGATCTTGAAGATGAGAGATGCTTGATGATTTGTGATGTGACC 183
 QY 314 LysAspGluLysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValIle 333
 DB 184 AAAGATGAAGAACCAAGTGTGATGCTGAGCACTCATTTGTAGAAGCAAGGCTGTG 243
 QY 334 AlaAspGlyHisIleSerTrpAlaGlyGluAlaPheSerArgPheThr---GluLysGln 352
 DB 244 GCAGATGTCATATTCCTTGGGCGATGAGGCTTTTCAAACTCATGCTGACGAGGTTT 303
 QY 353 LeuHisArgTrpSerSerLeuPheTrpCysTrpArgLeuPheLeuIleLysLeuTrpAsn 372
 DB 304 GCCCAAGCACACGACCATATGCAAGGCTTGGAATTTATCATATGATGAAGCTGTGCAAC 363
 QY 373 HisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArg 392
 DB 364 CATGCGCTGTTGATGCGCGGTACATTTAACTATTAACCTTAATATTATAGAGCAGTTCAA 423
 QY 393 AsnSer---SerAspThrThr 398
 DB 424 AACCAATACAGTACTACT 444
 RESULT 4
 BO505017 769 bp mRNA linear EST 22-JUL-2002
 LOCUS EST612432 Generation of a set of potato cDNA clones for microarray
 DEFINITION analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
 3' end, mRNA sequence.
 ACCESSION BO505017
 VERSION BO505017.2 GI:21920982
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 1 (bases 1 to 769)
 Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C.,
 Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
 Karamycheva, S.A.

QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 172 CACATCTCTACTTTCTTAAGAGATGTTGGCTTTAAATTAAGCAAGCAAAAAAGG 231
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 DB 232 AGGTGGCCAGCA---GGAATGTGTGTTTCAACTATAGGAGCTGTACAAATCCGCTCGA 288
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysLysPhe 100
 DB 289 AAGACTGAAGTGAAGTGAAGACTTTCTTCTGATTTCTGATTCAGTCCAGTCCGAGCTTT 348
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluLysLeu 120
 DB 349 AAGGTGTCGATTCATCTTTGTCATCAGATGATTCATTCATTCATTCAGTTCGGGTC 408
 QY 121 PheGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 409 ACCAATGATTCACCAAGCAGTGAACGCTCTCTGTAATAAACCTGAC-----ATATTGAGATCA 462
 QY 141 GluGlySerAspAsp-----AspLysPheGluProPheSerLeuCysSerLysPro 157
 DB 463 GAGATGTGCTGATGAGATGATTCACCAATCACAACCTCTCTCTCTCTCTCTCTCTCTCT 522
 QY 158 ArgLysArgArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
 DB 523 CGAAGCGCTAGACCAAGAGCGCTCGCTCAAAATGGGAAGCATGCCAATGTAATTTCTG 582
 QY 178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIle----- 192
 DB 583 AAGTGGATTCACCA-----GAAAGCATACGAATGGGCTCTACAGAAAGACAAC 633
 QY 193 -----ThrLeu 194
 DB 634 CATATCTCTCTCAGCAAAAGGAGCAATATGTCACAGATCTCGTAGAGAGAAATTTG 693
 QY 195 LeuAsnAspGlyAsnArg-----GlyLeuGlyTyrProGluAlaThrGluLeu 210
 DB 694 CAGATGAGGGAATAGTTGGGGTAAATTTGGCTGATTCAGTACAGAGCCATGATTTAC 753
 QY 211 AlaGly-GlnPheGluMetThrSerAsnIleProProAla 223
 DB 754 GCGGACATGTGTGAATCTAGTTTAAACATTTCCAGTCT 793
 RESULT 6
 LOCUS AJ469021 480 bp mRNA linear EST 24-MAY-2002
 DEFINITION AJ469021 S00008 Hordeum vulgare cDNA clone S000080053C04F1, mRNA
 sequence.
 ACCESSION AJ469021
 VERSION AJ469021
 KEYWORDS EST.
 SOURCE
 ORGANISM
 Hordeum vulgare.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 Triticeae; Hordeum.
 1 (bases 1 to 480)
 Saren,A.-M., Tanskanen,J., Paulin,L. and Schulman,A.H.
 Barley EST's
 Unpublished (2002)
 Contact: Schulman AH
 Institute of Biotechnology
 University of Helsinki
 P.O.Box 36 (Vilkkinkaari 6A), University of Helsinki FIN-00014,
 Finland.
 FEATURES
 source 1..480
 Location/Qualifiers
 /organism="Hordeum vulgare"
 /db_xref="taxon:4513"
 /clone="S000080053C04F1"
 /clone_lib="S00008"
 /tissue_type="Callus"

BASE COUNT 138 a 95 c 111 g 136 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 5,31e-51 Length: 480
 Score: 528.00 Matches: 95
 Percent Similarity: 79.14% Conservative: 15
 Best Local Similarity: 68.35% Mismatches: 29
 Query Match: 22.20% Indels: 0
 DB: 9 Gaps: 0
 US-09-890-220-2 (1-445) x AJ469021 (1-480)
 QY 257 GUAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
 DB 1 GATCCACAGAACCGAACACTCTGAGAAACGCTGTTTCCATTCACAGGCGACAG 60
 QY 277 PrometAlaLeuGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 296
 DB 61 CCAATGGCAGCTGGAAACAGTTTCTCGGATCGTATAGTAAGATGAAGTATGATGAC 120
 QY 297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
 DB 121 ATCCGAGATTTTGAAGATAGACGATGCTGATGATTTGTTGATGATCAGCAATGATGAG 180
 QY 317 LysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
 DB 181 AAATCTATTATGCAATATGCGAATTCATTTGCGAAACAAAGGCTCTACCTATGAT 240
 QY 337 HisIleSerTrpAlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyr 356
 DB 241 CATATTCCTGTGGCGCTGTGAGGATTCCTCCGCGCTTACATGAAACATCTTACAGAT 300
 QY 357 SerSerLeuPheTyrCysTrpArgLeuPheLeuIleLysLeuTyrAsnHisGlyLeuVal 376
 DB 301 CCTCTCTACATATGAGAGTGGCGTTCCTTATGATTAACCTGGAACACACAGTCTATTA 360
 QY 377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 395
 DB 361 GATGCCCGCTGCATGATGCTGCGGACCAATTCCTTCAAGGCTACCAAAATGAAGC 417
 RESULT 7
 LOCUS B0740672 561 bp mRNA linear EST 17-JUL-2002
 DEFINITION B0740672 S00005 Y1 Gm-c1076 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-c1076-3274 5' similar to TR:Q92NT9 Q92NT9
 FERTILIZATION-INDEPENDENT SEED 2 PROTEIN. ; mRNA sequence.
 ACCESSION B0740672
 VERSION B0740672
 KEYWORDS EST.
 SOURCE
 ORGANISM
 soybean.
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 1 (bases 1 to 561)
 Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna
 A., Bolla,B., Maria,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 R., Waterston,R. and Wilson,R.
 Public Soybean EST Project
 Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

This clone is available through: Resgen, Invitrogen Corp. 2130 South Memorial Parkway Huntsville, AL 35801 For further information call: (800)-533-4363 or contact: ccr@resgen.com web site: www.resgen.com

Seq primer: -40RP from G1bco
High quality sequence stop: 426.

FEATURES

source

Location/Qualifiers

1..561
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="SOYBEAN CLONE ID: Gm-cl076-3274"
/clone_1lb="Gm-cl076"
/tissue_type="wounded cotyledons"
/dev_stage="11 day old seedlings"
/lab_host="DH10B"

/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 11 day old seedlings treated with that were treated with 2 ugs/ml of a crude glucan elicitor preparation isolated from the mycelial walls of Phytophthora sojae. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. Plant material was provided by Michael G. Hahn (Complex Carbohydrate Research Center, University of Georgia) and the library was constructed by Anu Khanna (Lila Vodka lab University of Illinois)."

BASE COUNT 166 a 91 c 123 g 181 t
ORIGIN

Alignment Scores:

Pred. No.: 7.41e-49 Length: 561
Score: 510.50 Matches: 93
Percent Similarity: 77.78% Conservative: 19
Best Local Similarity: 64.58% Mismatches: 27
Query Match: 21.47% Indels: 5
DB: 14 Gaps: 1

US-09-890-220-2 (1-445) x B0740672 (1-561).

OY 280 Leunglunlvalmetseraspargaspsercunaspqlunlaspaspaspvalalaasp 299
DB 2 CTAGAACAAAGTGTATACAGCCGTGATGAGAACGAGTGTGACGACATTCGACAGAT 61
OY 300 Pheglunasparglmetleuasppsphevalaspvalanulaspaspaspaspaspasp 319
DB 62 CTTGAGATAGAGAGATGCTTGCAGATTTGTGATCTTCCAAAGATGAAAAACAGCTC 121
OY 320 Methlslstrpasnserphevalarglsglnarvallealaspaglyhslsleser 339
DB 122 ANGCACTCTGGAACCTTTTATAGAGCAAGAGGTGTCAGAGATGTCATGTCG 181
OY 340 TTPAlacysglualapheserfarphefysglunlunlunhsarqfyserserleu 359
DB 182 TGGGCGCTGAGGCTTTTCCACGCTTCATGAAAAGAGTGCATTCATCCACCTTTA 241
OY 360 Pheftrpcystpfpaglleupheleuileylsleutrpasnhsiglyleuvalaspserala 379
DB 242 TTTTGGGTGGAGGTATTCATGATCAACTTGGATCATGATGATGCTTCTTGTGATGCTCTG 301
OY 380 ThrileasnAsnCyasnThrileleuglunsnCyasrfgasnseraseraspftrhrthr 399
DB 302 ACAATGAACAACCTAGCATAGTATAGTACGAATACAGGATCGGCTACAATA 361
OY 400 ThrAsnAsnAsnSerValaspargproseraspserasnhrAsnAsnAsnAnile 419
DB 362 AAAAAT-----TGAAGACCATTAAGGAACATGACTCAACAAACAGATATA 406

OY 420 ValaspHsipro 423
DB 407 CAGTCTTCTCCG 418

FEATURES

source

Location/Qualifiers

1..600
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE3451_E10_119"
/clone_1lb="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

DEFINITION B1479743 600 bp mRNA linear EST 28-AUG-2001
LOCUS B1479743
DEFINITION WHE3451_E10_119S Wheat pre-anthesis spike cDNA library Triticum
ACCESSION B1479743
KEYWORDS B1479743.1 GI:15323794
SOURCE EST
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae
; Triticaceae; Triticum.
1 (bases 1 to 600)
Anderson,O.D., Chao,S., Choi,D.W., Close,T.J., Fenton,R.D., Han,P.S., Hsia,C.C., Kang,Y., Lazo,G.R., Miller,R., Rausch,C.J., Seaton,C.L. and Tong,J.C.
The structure and function of the expressed portion of the wheat genomes - Pre-anthesis spike cDNA library
Unpublished (2000)
Contact: Olin Anderson
US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
800 Buchanan Street, Albany, CA 94710, USA
Tel: 5105959773
Fax: 5105959818
Email: oanderson@wr.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.

FEATURES

source

Location/Qualifiers

1..600
/organism="Triticum aestivum"
/cultivar="Chinese Spring"
/db_xref="taxon:4565"
/clone="WHE3451_E10_119"
/clone_1lb="Wheat pre-anthesis spike cDNA library"
/tissue_type="Spike before anthesis"
/dev_stage="Adult plant"
/lab_host="E. coli SOLR"
/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI; Site 2: XhoI; Plants were grown in the greenhouse. Whole spike with awns trimmed, white, green and yellow anther were collected and total RNA, and poly(A) RNA were prepared, a cDNA library was made, and the cDNA clones were in vivo excised to give plasmid phagemids in the T7 Close lab (Choi, Close, Fenton) at the University of California, Riverside. Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT

165 a 126 c 143 g 166 t

Alignment Scores:

Pred. No.: 1.23e-48 Length: 600
Score: 509.00 Matches: 102
Percent Similarity: 62.18% Conservative: 18
Best Local Similarity: 52.85% Mismatches: 44
Query Match: 21.40% Indels: 2
DB: 13 Gaps: 2

US-09-890-220-2 (1-445) x B1479743 (1-600)

OY 237 LeuthregerlunlvalaProalatrThrlyfysleuSerlAglylrgser 256
DB 21 CTTACACACGACAGTACTAGAGTTTGGGAAGACAGAACTATCTGCGAGGAGT 80
OY 257 Glunlalarger----- 260

(pbluescript II SK(+)) that has been digested with EcoRI and XhoI, and phosphorylated by StrataGene). Both the white and blue colonies appear to contain recombinant plasmids with cDNA inserts, based on size (n=25). This library was constructed by Dr. Paul Keim and Dr. Virginia Coryell."

BASE COUNT 133 a 86 c 107 g 148 t
ORIGIN

Alignment Scores:

Pred. No.: 1.14e-42 Length: 474
Score: 456.50 Matches: 89
Percent Similarity: 78.36% Conservative: 16
Best Local Similarity: 66.42% Mismatches: 28
Query Match: 19.20% Indels: 1
DB: 10 Gaps: 1

US-09-890-220-2 (1-445) x AM234600 (1-474)

OY 1 MetCysArgGlnAsnGlyArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
|||||
DB 43 ATGTCCCGCAAAATTCCTCCGTACACCATGCCGTGAAGAAATTCACCTGATGAG 102
OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
:::|||||
DB 103 AGCTTTAAATTTATTGCAAGCCTGTGAACGTGTACAAATATTCTACCGCGCTCTT 162
OY 41 GlyAsnProSerPheLeuProArgGlySerLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
|||||
DB 163 CAAATTCCTTTCTTGTAGAGATGATGCTTTTAAATTAAGAGAGGCGTAAAGG 222
OY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnThrLeuGln 80
:::|||||
DB 223 AGCTTGAGAGCA---GGAATGTGTATTTCAATTTAGGAGTACATACACATTCTTCGG 279
OY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
|||||
DB 280 AAAACTGAGTACGACGAGACACTTTCTGTCCGTTTCTGATGACATGAGTGAAGCTTT 339
OY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
|||||
DB 340 AAGGCTTTCGATTCATCTTTGTTGATCACAATGATTCATTCACCTTGAGTTTGGGTT 399
OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
|||||
DB 400 ACTGAAGATTACCAAGCAGTGAATGCTCCGCGAATAATTAA 441

RESULT 12

LOCUS BG580495

DEFINITION EST1482220 GVN Medicago truncatula cDNA clone pGVN-57B2 5' end, mRNA

ACCESSION BG580495

VERSION BG580495.1

KEYWORDS

SOURCE

ORGANISM

Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 467)
Fedorova,M., Plešková,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town
C.D., Van Aken,S., Utecher,T., Cho,J. and Fraser,C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula, 2001
Unpublished (2001)
Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651 649-5058

Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M381546e TIGR sequence name:
MTCOT01TK More information is available at: <http://www.medicago.org>
Seq primer: SKmod (CTA GAA CTA gta gAT CC).
Location/Qualifiers

FEATURES

source

1..467

/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pGVN-57B2"
/clone_11b="GVN"

/tissue_type="N2-fixing root nodules"

/dev_stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"

/lab_host="E. coli strain XL0R"

/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
stratagene and packaged using GigaPack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XL0R cells."

BASE COUNT 136 a 87 c 95 g 149 t
ORIGIN

Alignment Scores:

Pred. No.: 6.39e-42 Length: 467
Score: 450.00 Matches: 91
Percent Similarity: 70.32% Conservative: 18
Best Local Similarity: 58.71% Mismatches: 40
Query Match: 18.92% Indels: 6
DB: 12 Gaps: 3

US-09-890-220-2 (1-445) x BG580495 (1-467)

OY 17 SerThrAspGluAsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHis 36
:::|||||
DB 3 ACAGCTGATGAAAGCTTTTGTATTTATTCGCAACCTGTTGAATTTCTACATTTCTGAT 62
OY 37 LeuArgSerLeuGlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAla 56
|||||
DB 63 CGTCGTTCTTCAACAATCTTCTTTTGAAGATGTTGCGGTATATAAATAAGCA 122
OY 57 LysArgLysArgLysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsn 76
|||||
DB 123 AAGCAAAAAGAGAGGTTGCGAGCA---GGAATGTGTGTTTCAACTATAGGACCTGTAC 179
OY 77 AsnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPheCysSerMetLeu 96
|||||
DB 180 AATGCGCTTCGAAAGACGTAAGTGAAGTGAAGCTTTCTGTCATTTGCTGATGAG 239
OY 97 CysGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPhe 116
|||||
DB 240 TGTGCGACTTTAAAGGTTTGGATTCATCTTTTGCATGACATGATCTATTCACCTTT 299
OY 117 GluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
|||||
DB 300 GAGTCTGAGGTACCAATGATTACCAAGCAGTGAAGCTCTGTAATAACTGAC----- 353
OY 137 IlePheGluGluGlySerAspAsp-----AspLysPheGluProPheSerLeu 153
|||||
DB 354 ATATTGATGATGAGAAATGTTGCTGATGAGTAATTCACCAATCAACCAACCTTCTTTC 413
OY 154 CysSerLysProArgLysArgArgGlnArgLysGlyLysArgAsn 168
|||||
DB 414 TGTTCAGATCTCGAAAGCGTAGACAGAGGCTCCGTCAAAAT 458

RESULT 13

LOCUS BG314044

DEFINITION WHE2467_D04_G07Z5 Trifolium monocoecum early reproductive apex cDNA

library Trilicium monoccoccum cDNA clone WHE2467_D04_G07, mRNA sequence.
 ACCESSION BG314044
 VERSION BG314044.1 GI:13115847
 KEYWORDS EST.
 SOURCE Trilicium monoccoccum.
 ORGANISM Trilicium monoccoccum.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidaeae; 1 (bases 1 to 405)
 REFERENCE Anderson, O.D., Chao, S., Dubcovsky, J., Echenique, V., Han, P.S., Hsia, C.C., Kang, Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L., Stamova, B. and Tong, J.C.
 The structure and function of the expressed portion of the wheat genomes - Early reproductive apex cDNA library from Trilicium monoccoccum
 TITLE Unpublished (2001)
 JOURNAL Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105555773
 Fax: 5105555818
 Email: oanderson@w.usda.gov
 Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
 Seq primer: Stratagene SK primer.
 FEATURES
 source
 1.405
 /organism="Trilicium monoccoccum"
 /cultivar="DV92"
 /db_xref="taxon:4568"
 /clone="WHE2467_D04_G07"
 /clone_lib="Trilicium monoccoccum early reproductive apex cDNA library"
 /tissue_type="Early reproductive apex"
 /dev_stage="Seven week-old plants"
 /lab_host="E. coli XL0R"
 /note="Vector: lambda Uni-ZAP XR, excised phagemid; Site 1: EcoRI, Site 2: XhoI, the tissue, total RNA, and poly(A) RNA were prepared from apex at terminal-spikelet stage during transition from vegetative state to flower state, a cDNA library was made, and the cDNA clones were in vivo excised at the University of California, Davis (V. Echenique, B. Stamova, J. Dubcovsky). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."
 BASE COUNT 105 a 85 c 104 g 111 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8.8e-42 Length: 405
 Score: 448.00 Matches: 78
 Percent Similarity: 81.58% Conservative: 15
 Best Local Similarity: 68.42% Mismatches: 21
 Query Match: 18.84% Indels: 0
 Gaps: 0
 DB: 12
 US-09-890-220-2 (1-445) x BG314044 (1-405)
 QY 276 GlnPromeTalaLeuGlInValMetSerAspArgSPSerGluAspGluValAspASP 295
 |||||||
 DB 4 CAGCCAAATGCGAGCAAGATTGTCAGATATACGCGGATGAGTGAAGTTGATGAT 63
 |||||||
 QY 296 AspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAspASP 315
 |||||||
 DB 64 GACATTGCCGACTTGCAGATAGACGATGCTTCTGATTTCTGATGTCAGGAAGAT 123
 |||||||
 QY 316 GAluysGlnPheMetHisLeuTrpAsnSerPheValArgLysGlnArgValAlaAsp 335
 |||||||
 DB 124 GAGAAAGCTTATATATCATATGTGCAATTCATTATTCGAAACAAAGGATGCTAGCTGAT 183
 |||||||

336 GlyHisIleSerTrrPalacysGluAlaPheSerArgPheTyrGluLysGluLeuHisArg 355
 |||||||
 DB 184 GGGCATATACCTTGAGCGCTGTGAGCGCTTCCCGGCTTATGACCGACGCTTGACAA 243
 |||||||
 QY 356 TyrSerSerLeuPheTrrPcysTrrPArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeu 375
 |||||||
 DB 244 AACCTCTCTGCTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
 |||||||
 QY 376 ValAspSerAlaThrIleAsnAsnCysAsnTrpIleLeuGlu 389
 |||||||
 DB 304 CTGATGCGCGCGCCGACATGACATGACATGACATGACATGACATGACATGACATGACATG 345
 |||||||
 RESULT 14
 BQ279630/c 662 bp mRNA linear EST 08-MAY-2002
 LOCUS 1091030A10.x2.1091 - Immature ear with common ESTs screened by
 DEFINITION Schmidt lab Zea mays cDNA, mRNA sequence.
 ACCESSION BQ279630.1 GI:20507433
 VERSION BQ279630.1
 KEYWORDS EST.
 SOURCE Zea mays.
 ORGANISM Zea mays.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoidae; Andropogoneae; Zea.
 1 (bases 1 to 662)
 REFERENCE Walbot, V.
 Zea mays
 Maize ESTs from various cDNA libraries sequenced at Stanford University
 JOURNAL Unpublished (1999)
 COMMENT Contact: Walbot V
 Department of Biological Sciences
 Stanford University
 855 California Ave, Palo Alto, CA 94304, USA
 Tel: 650 723 2227
 Fax: 650 725 8221
 Email: walbot@stanford.edu
 Plate: 1091030 row: A column: 10.
 FEATURES
 source
 1.662
 /organism="Zea mays"
 /cultivar="OH43"
 /db_xref="taxon:4577"
 /clone_lib="1091 - Immature ear with common ESTs screened by Schmidt lab"
 /tissue_type="Inflorescence meristem - floral organ primordia"
 /dev_stage="0.5 cm to 2 cm"
 /lab_host="Stratagene XL0R"
 /note="Organ: Immature ear; Vector: PAD-GAL4; Site 1: EcoRI; Site 2: XhoI; RNA from library 606 was filtered for common ESTs found in 606."
 BASE COUNT 208 a 132 c 111 g 211 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.76e-41 Length: 662
 Score: 444.50 Matches: 88
 Percent Similarity: 64.78% Conservative: 15
 Best Local Similarity: 55.35% Mismatches: 25
 Query Match: 18.69% Indels: 31
 Gaps: 2
 DB: 14
 US-09-890-220-2 (1-445) x BQ279630 (1-662)
 QY 237 LeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSer 256
 |||||||
 DB 622 CTTTCACACACACACAGTACAGTGTGGAGACAGACAGACATCTCT---GAGGATCT 566
 |||||||
 QY 257 GAluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValAla 276
 |||||||
 DB 565 GACCTTAGACATGCGGACACTCTGCAAAAACGACAGTCTTCCATCTTCACAGGCGCAG 506
 |||||||

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:23:13 ; Search time 23 Seconds

(without alignments)
802.477 Million cell updates/sec

Title: US-09-890-220-2

Sequence: 2378 1 MCRONCRKAKSSPEVISTDE.....INNNKNNVDKNNNSDKVIK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138.5	5.8	989	1	PRP3_DICDI
2	132	5.6	749	1	MAD1_YEAST
3	127.5	5.4	448	1	AAC2_DICDI
4	126.5	5.3	1584	1	KYK1_DICDI
5	123.5	5.2	1858	1	P3K2_DICDI
6	120	5.0	1000	1	S1S5_YEAST
7	120	5.0	1195	1	YK76_YEAST
8	119	5.0	1094	1	YB00_YEAST
9	117.5	4.9	988	1	OMB_DROME
10	117	4.9	2175	1	HMCU_DROME
11	114	4.8	964	1	YIN0_YEAST
12	112.5	4.7	2278	1	FAB1_YEAST
13	112	4.7	1314	1	SWI1_YEAST
14	111.5	4.7	735	1	C1GB_DICDI
15	111.5	4.7	858	1	CYAG_DICDI
16	108	4.5	1407	1	HSE_YEAST
17	107.5	4.5	833	1	PC11_PLAFA
18	107.5	4.5	2339	1	PC11_PLAFA
19	106.5	4.5	1585	1	HAAL_YEAST
20	106	4.5	779	1	SRP_DROME
21	106	4.5	779	1	SRP_DROME
22	105.5	4.4	558	1	YKXA_ASTLO
23	105	4.4	396	1	PMAR_YEAST
24	105	4.4	555	1	POST_XENLA
25	105	4.4	732	1	KHMB_DICDI
26	104.5	4.4	666	1	YEA7_YEAST
27	104.5	4.4	935	1	COPE_YEAST
28	104.5	4.4	1419	1	MDR_PLAFA
29	104	4.3	871	1	SC10_YEAST
30	103	4.3	432	1	AP2_ARATH
31	103	4.3	490	1	CAR3_DICDI
32	103	4.3	1085	1	CUT7_SCHRO
33	102.5	4.3	481	1	BIND_STRPU

34	102.5	4.3	1076	1	RPOR_ASTIO
35	102.5	4.3	1570	1	P3K1_DICDI
36	102.5	4.3	2733	1	RRPB_CWAS
37	101.5	4.3	647	1	KNRL_DROME
38	101.5	4.3	946	1	FTSR_CAMJE
39	101.5	4.3	1341	1	YL78_YEAST
40	101.5	4.3	2731	1	RRPB_CWMIH
41	101	4.2	534	1	GCR2_YEAST
42	101	4.2	537	1	ARE_PLAFA
43	101	4.2	1131	1	YAB9_YEAST
44	100.5	4.2	1188	1	OSH1_YEAST
45	100	4.2	284	1	SMX5_SCHMA

ALIGNMENTS

RESULT 1	ID	PRP3_DICDI	STANDARD:	PRT:	989 AA.
AC	P54637				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Protein-tyrosine phosphatase 3 (EC 3.1.3.48) (Protein-tyrosine-phosphatase 3)				
DE	phosphatase phosphohydrolase 3)				
GN	(PRP1 OR PRP3) AND (PTP2 OR PTP3)				
OS	Dicystostellum discoidium (Slime mold).				
OC	Eukaryota; Mycetozoa; Dictyostelida; Dictyostellum.				
OX	NCBI_TaxID=44689;				
RN	[1]				
RN	SEQUENCE FROM N.A.				
RC	STRAIN=AX3.				
RX	MEDLINE=96189126; PubMed=8628311;				
RA	Gamper M., Howard P.K., Hunter T., Firtel R.A.;				
RT	"Multiple roles of the novel protein tyrosine phosphatase PRP3 during Dictyostellum growth and development."				
RL	Mol. Cell. Biol. 16:2431-2444(1996).				
CC	- FUNCTION: SEEMS TO DEPHOSPHORYLATE A PROTEIN OF 130 kDa (P130).				
CC	- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O -> protein-tyrosine + phosphate.				
CC	- SUBCELLULAR LOCATION: Cytoplasmic.				
CC	- TISSUE SPECIFICITY: IN THE ANTERIOR-LIKE AND PRESTALK CELL TYPES.				
CC	- DEVELOPMENTAL STAGE: EXPRESSED AT MODERATE LEVELS DURING GROWTH AND DEVELOPMENT.				
CC	- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb.ch/announce/ or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: U38197; AAC47041.1; -				
DR	HSSP: 006124; 2SRP.				
DR	DiclydD: DD01111; pepC1.				
DR	DiclydD: DD07777; pepC2.				
DR	InterPro: IPR000387; Tyr_phosphatase.				
DR	InterPro: IPR000242; Tyr_PP.				
DR	Pfam: PF00102; Y.phosphatase.1.				
DR	PRINTS: PR00700; KRYPPHRTASE.				
DR	SMART: SM00194; PTPc.1.				
DR	PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.				
DR	PROSITE: PS50056; TYR_PHOSPHATASE_2; 1.				
DR	PROSITE: PS50055; TYR_PHOSPHATASE_PTP; 1.				
KW	Hydrolase.				
FT	ACT_SITE 649				
FT	DOMAIN 460				
FT	DOMAIN 64				
FT	DOMAIN 71				
FT	DOMAIN 109				
FT	DOMAIN 118				

BY SIMILARITY.
PROTEIN-TYROSINE PHOSPHATASE.
POLY-ASN.
POLY-ASN.

FT	DOMAIN	137	190	POLY-ASN.
FT	DOMAIN	249	257	POLY-SER.
FT	DOMAIN	258	265	POLY-THR.
FT	DOMAIN	286	289	POLY-ASN.
FT	DOMAIN	366	371	POLY-SER.
FT	DOMAIN	787	790	POLY-GLN.
FT	DOMAIN	834	839	POLY-GLN.
FT	DOMAIN	883	892	POLY-GLN.
FT	DOMAIN	906	914	POLY-ASN.
FT	DOMAIN	943	963	POLY-ASN.
SO	SEQUENCE	969 AA:	109995 MW:	95711055F8097AAF CRC64:

Query Match	5.88;	Score 138.5;	DB 1;	Length 989;
Best Local Similarity	19.88;	Pred. No. 0.0064;		
Matches 72; Conservative	51;	Mismatches 140;	Indels 101;	Gaps 12

```

0Y      129  V$V$KLS$F$EE$G$S$D$D$F$E$F$S$C$S$P$R$R$R$R$G$G$N$N$R$R$K$V$C$F$P$D$S$P$T$N$G 188
Db      660  I$T$IN$M$A$K$E$R$G$N$D$P$S$O$M$IN$I$S$D$V$L$E$R$R$O$R$G$W$V$O$-----L$D$Y$I$F$E$V 710
0Y      189  E$N$G$I$T$L$D$N$R$G$I$G$P$E$A$T$E$F$L$A$G$O$E$M$T$-----N$I$P$A$I$-----H$S$D$A$C$A$K$Y 236
Db      711  I$N$D$V$-----L$T$D$W$G$I$R$E$L$S$P$S$K$R$C$E$M$I$K$T$P$M$P$R$I$D$I$S$I$P$P$L$T$P$K$D$O$S$T$P$S$T$D$M 769
0Y      237  L$T$B$A$V$P$A$P$K$T$R$K$T$S$A$E$R$S$E$-----A$R$S$H$I$L$L$-----O$K$O$Y$H$S$H$R$O$P$-M 278
Db      770  A$S$I$T$I$T$O$M$Q$T$K$P$P$Q$O$O$N$P$P$S$K$S$I$K$P$S$P$L$N$T$N$I$S$T$K$P$Q$O$P$H$O$P$O$L 829
0Y      279  A$L$E$Q$V$M$S$D$R$E$D$E$V$D$D$-----V$A$D$E$R$O$M$L$D$F$V$Y$N$K$D$E$K$O$F$H$I$M$N 324
Db      830  D$L$N$Q$O$Q$O$Q$S$Q$O$L$D$N$D$P$L$M$M$S$N$S$I$K$F$P$Y$V$S$I$S$C$H$E$D$S$K$N$D$N$N$Q$----- 883
0Y      325  S$F$A$R$Q$V$Y$A$D$G$I$T$S$A$C$A$F$S$F$E$F$E$K$E$L$H$R$S$I$F$W$C$R$L$F$I$K$N$H$G$L$V$A$T$I$N$C 384
Db      884  ---Q$O$Q$O$Q$O$Q$K$N$N$O$C$G$F$S$H$-----L$N$N 908
0Y      385  N$T$I$E$N$C$R$-----N$S$D$T$T$Y$N$N$N$S$V$D$P$S$D$N$T$N$N$N$I$V$D$H$P$N$D$I$N$K$N$Y$D$K$N$- 438
Db      909  N$N$D$N$N$O$S$S$G$G$F$N$G$F$L$E$N$S$N$N$S$S$T$S$E$C$S$N$N$K$N$N$N$N$N$N$N$N$N$N$N$K$N$S$D$N$G 968
0Y      439  S$R$D$K 442
Db      969  T$K$D$K 972

```

ID	MAD1_YEAST	STANDARD:	PRT;	749 AA.
AC	P40957;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Spindle assembly checkpoint component MAD1 (Mitotic MAD1 protein).			
GN	MAD1 OR IGLO86M.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetaceae; Saccharomycetaceae; Saccharomycetes.			
OX	NCBI_TaxID:4932;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	STRAIN-W303;			
RC	MEDLINE=96042315; PubMed=7593191;			
RA	Hardwick K.G., Murray A.W.;			
RT	"Mad1, a phosphoprotein component of the spindle assembly checkpoint			
RL	in budding yeast.";			
RL	J. Cell Biol. 131:709-720(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-S288C;			
RC	MEDLINE=97455481; PubMed=9290212;			
RA	Rieger M., Bruckner M., Schaefer M., Mueller-Auer S.;			
RT	"Sequence analysis of 203 kilobases from Saccharomyces cerevisiae			
RT	chromosome VII.";			

RL Yeast 13:1077-1090(1997).
 RN [3]
 RP INTERACTIONS.
 RX MEDLINE-98128031; PubMed-9461437;
 RA Hwang L.H., Lau L.F., Smith D.L., Mistrot C.A., Hardwick K.G.,
 RA Hwang E.S., Amon A.W., Murray A.W.;
 RT "Budding yeast Cdc20: a target of the spindle checkpoint.";
 RL Science 279:1041-1044(1998).
 CC -1- FUNCTION: CENTRAL COMPONENT OF THE SPINDLE ASSEMBLY CHECKPOINT.
 CC -1- SUBUNIT: THE SPINDLE CHECKPOINT COMPLEX IS COMPOSED OF MAD1, MAD2
 CC AND MAD3. IT INTERACTS WITH CDC20.
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- PPM: BECOMES HYPERPHOSPHORYLATED WHEN WILD-TYPE CELLS ARE ARRESTED
 CC IN MITOSIS.
 CC -1- SIMILARITY: SOME. TO S.POMBE SPCC305.04C.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR	EMBL	U14632	AA91620.1	-
DR	EMBL	Z72608	CAB96791.1	-
DR	PIR	S48246	S48246	
DR	SGD	S0003054	MAD1	
KW	Mitos: Col	Colled coil	Nuclear protein; Phosphorylation.	
FT	DOMAIN	57	321	COILED COIL (POTENTIAL).
FT	DOMAIN	253	324	COILED COIL (POTENTIAL).
FT	DOMAIN	390	656	COILED COIL (POTENTIAL).
FT	DOMAIN	323	329	COILED COIL (POTENTIAL).
FT	DOMAIN	330	347	POLY-ASP.
FT	DOMAIN	355	363	POLY-ASN.
FT	DOMAIN	601	604	POLY-LEU.
SO	SEQUENCE	749 AA	87651 MW	312DD1BE2A41610D CRC64

[illegible]

```

ID AAC2_DICDI STANDARD; PRT; 448 AA.
AC P14196;
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE AAC-rich mRNA clone AAC11 protein (Fragment).
OS Dictyostellum discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=90066348; PubMed=2511421;
RA Shaw D.R., Richter H., Giorda R., Ohmachi T., Ennis H.L.;
RT "Nucleotide sequences of Dictyostellum discoideum developmentally
regulated cDNAs rich in (AAC) imply proteins that contain clusters of
asparagine, glutamine, or threonine."
RL Mol. Genet. 218:453-459(1989).
CC -1- DEVELOPMENTAL STAGE: THE CONCENTRATION OF AAC-RICH MRNAS IS LOW
CC IN DORMANT SPORES AND GROWING CELLS, BUT INCREASES DURING
CC SPORE-GERMINATION AND MULTICELLULAR DEVELOPMENT.
CC -1- MISCELLANEOUS: SEVERAL PROTEINS DERIVE FROM AAC-RICH MRNA, WHICH,
CC DUE TO A FRAMESHIFT ALSO HAVE ACA AND CAA CODONS AND THUS ARE
CC ASN-, THR- OR GLN-RICH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X16522; CAA34529.1; -
DR PIR: S05355; S05355.
DR DictyDb: DD05006; -.
DR InterPro: IPR000637; AT_hook.
DR InterPro: IPR000116; Highmobility_IV.
DR Pfam: PF02178; AT_hook; 4.
DR ProDom: PD005593; Highmobility_IV; 1.
DR SMART: SM00384; AT_hook; 4.
DR Repeat.
KW NON_TER
FT DOMAIN 1 25 GLN-RICH.
FT DOMAIN 115 144 ASN-RICH.
FT DOMAIN 268 384 ASN-RICH.
SQ SEQUENCE 448 AA; 48636 MW; D8FD80D910D99817 CRC64;

Query Match 5.4%; Score 127.5; DB 1; Length 448;
Best Local Similarity 35.7%; Pred. No. 0.016; 24; Indels 1; Gaps 1;
Matches 25; Conservative 20; Mismatches 24;

OY 372 NHGLVDSATINNCATILECRNSDPTTTNNNSVDRSDSTNNNNIYDHPNDINKNN 431
DB 273 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 331
OY 432 VDNKDNNSRD 441
DB 332 INNANNNTNN 341

RESULT 4
KYK1_DICDI STANDARD; PRT; 1584 AA.
AC P18160;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Non-receptor tyrosine kinase spore lysis A (EC 2.7.1.112) (Tyrosine-
protein kinase 1).
GN PYKA OR SPLA OR DPYK1.
OS Dictyostellum discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostellum.
OX NCBI_TaxID=44689;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH10.
RX MEDLINE=97053827; PubMed=8898241;
RA Nuckolls G.H., Oshero N., Loomis W.F., Spudich J.A.;
RT "The Dictyostellum dual-specificity kinase splA is essential for
spore differentiation."
RL Development 122:3295-3305(1996).
RN [2]
RP SEQUENCE OF 1248-1584 FROM N.A.
RX MEDLINE=90287147; PubMed=1972546;
RA Tan J.L., Spudich J.A.;
RT "Developmentally regulated protein-tyrosine kinase genes in
Dictyostellum discoideum."
RL Mol. Cell. Biol. 10:3578-3583(1990).
CC -1- FUNCTION: ESSENTIAL FOR SPORE DIFFERENTIATION.
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
CC tyrosine phosphate.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT WITH A PEAK
CC DURING THE MOUND STAGE OF MORPHOGENESIS.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES.
CC -1- SIMILARITY: CONTAINS 1 SAM DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: U32174; AAB41125.1; -
DR PIR: M33785; AAA33202.1; -.
DR DictyDb: DD03010; PYKA.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR003877; SPRY_domain.
DR InterPro: IPR003877; SPRY_receptor.
DR InterPro: IPR004040; STY_pkinase.
DR InterPro: IPR001245; TYR_pkinase.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00536; SAM; 1.
DR Pfam: PF00622; SPRY; 3.
DR ProDom: PD000001; Euk_pkinase; 1.
DR SMART: SM00454; SAM; 1.
DR SMART: SM00449; SPRY; 3.
DR ProSITE: PS00221; STYKC; 1.
DR ProSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR ProSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR ProSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR ProSITE: PS50105; SAM_DOMAIN; 1.
KW Transferase; Tyrosine-protein kinase; ATP-binding; Phosphorylation.
FT DOMAIN 908 972 SAM.
FT DOMAIN 403 420 POLY-ASN.
FT DOMAIN 428 435 POLY-THR.
FT DOMAIN 449 480 POLY-ASN.
FT DOMAIN 483 491 POLY-ASN.
FT DOMAIN 494 508 POLY-ASN.
FT DOMAIN 512 532 POLY-ASN.
FT DOMAIN 532 600 POLY-ASN.
FT DOMAIN 600 808 POLY-PHE.
FT DOMAIN 808 811 POLY-SER.
FT DOMAIN 1026 1029 POLY-SER.
FT DOMAIN 1195 1210 POLY-ASN.
FT DOMAIN 1215 1220 POLY-GLN.
FT DOMAIN 1224 1233 POLY-GLN.
FT DOMAIN 1266 1274 POLY-PRO.
FT DOMAIN 1289 1561 PROTEIN_KINASE.
FT NP_BIND 1295 1303 ATP (BY SIMILARITY).
FT BINDING 1316 1316 ATP (BY SIMILARITY).
FT ACT_SITE 1417 1417 BY SIMILARITY.
FT ACT_SITE 1417 1417 D -> R (IN REF. 2).
FT CONFLICT 1248 1248 V -> L (IN REF. 2).
FT CONFLICT 1435 1435

```

SEQUENCE 1584 AA: 174304 MW: 5D1589458DBE01E3 CRC64;

Query Match 5.3%; Score 126.5; DB 1; Length 1584;
Best Local Similarity 22.2%; Pred. No. 0.099; Mismatches 100; Indels 133; Gaps 15;
Matches 77; Conservative 37;

DB 149 EEPFSCSKPRKRORGGNNTRRLKVCFLPLDPSLNTGTEGTLINDGNGRLGYPEAT 208
DB 210 EEEFECRWSEKKKNH-GKHVYVRNRTAFLLPDSPTDITG---GVAAQDPFEGECY--- 261
DB 209 ELAOGFEMTSNIPPAIHAHSSDAGAKVLTSEAVVPATKRLKLSERSEARSHLLQKRO 268
DB 262 ----FE-----VIIDQDKGOLSIGLANLEPFP-----YHGVMPRS 295
DB 269 F-YHSHRYQPMALQVMSDRSEDEVDVADFEEDROMLDFVDVKNDEKQFMHLMNSFV 327
DB 296 YGYH-----NDGGRKFRMRRE---PGVNEGES-----YGSY 324
DB 328 KRQRYADGHISMACE-----AFSREYKELHRYSSLFMCWRLF----- 366
DB 325 KKQDIIGCG-LSFTSREIFETKNGMYLGTAFSANY-----GVFYDPAFNPGISIT 375
DB 367 -----LILKMHGLVDSATI-----NNCNTILENCSSDPT 398
DB 376 GVFGPPFKFSQVTLMLKN---VNSTSLVPGNKNNNNNNNNNNNNNNITIGKLTITTTT 432
DB 399 TT-----NNNSVDRPSDSMTNNNTVDHPNDINKKNVNDKDNNS 439
DB 433 TSTPSSINNEDISSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNS 479

RESULT 5

P342.DICDI STANDARD: PRT: 1858 AA.

AC P34674: 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Phosphatidylinositol 3-kinase 2 (EC 2.7.1.137) (PI3-kinase)
DE (PtdIns-3-kinase) (PI3K).
GN PI3K OR PI3K.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelidae; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX3;
RA MEDLINE=96009592; PubMed=7565716;
RA Zhou K., Takagawa K., Emr S.D., Firtel R.A.;
RT "A phosphatidylinositol (PI) kinase gene family in Dictyostelium
discoideum: Biological roles of putative mammalian p110 and yeast
Vps4p PI 3-kinase homologs during growth and development.";
RL Mol. Cell. Biol. 15:5645-5656(1995).
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidyl-1D-myo-inositol = ADP +
1-phosphatidyl-1D-myo-inositol 3-phosphate.
CC -1- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U23477; AAA85722.1;
DR Dictydb: DD01100; PI3K.
DR InterPro: IPR002420; PI3K_C2.
DR InterPro: IPR003441; PI3K_Cas_Bind.
DR InterPro: IPR01263; PI3K.
DR InterPro: IPR00403; PI3K_P14_Kinase.
DR Pfam: PF00454; PI3_P14_Kinase; 1.
DR Pfam: PF00613; PI3Ka; 1.

DR Pfam: PF00792; PI3K_C2; 1.
DR Pfam: PF00794; PI3K_C2; 1.
DR SMART: SM00142; PI3K_C2; 1.
DR SMART: SM00144; PI3K_C2; 1.
DR SMART: SM00145; PI3Ka; 1.
DR SMART: SM00146; PI3K; 1.
DR PROSITE: PS00915; PI3_4_KINASE_1; 1.
DR PROSITE: PS00916; PI3_4_KINASE_2; 1.
DR PROSITE: PS00917; PI3_4_KINASE_3; 1.
DR PROSITE: PS00918; PI3_4_KINASE_4; 1.
KW Transferase; Kinase; Multigene family.
FT DOMAIN 34 40 POLY-SER.
FT DOMAIN 166 172 POLY-SER.
FT DOMAIN 185 226 POLY-ASN.
FT DOMAIN 227 235 POLY-THR.
FT DOMAIN 246 253 POLY-SER.
FT DOMAIN 259 268 POLY-ASN.
FT DOMAIN 294 303 POLY-THR.
FT DOMAIN 361 364 POLY-SER.
FT DOMAIN 369 384 POLY-SER.
FT DOMAIN 425 429 POLY-GLN.
FT DOMAIN 439 444 POLY-THR.
FT DOMAIN 445 454 POLY-SER.
FT DOMAIN 562 570 POLY-GLY.
FT DOMAIN 715 727 POLY-THR.
FT DOMAIN 982 990 POLY-GLN.
FT DOMAIN 1015 1049 POLY-ASN.
FT DOMAIN 1598 1858 POLY-PI3K.
SQ SEQUENCE 1858 AA: 203945 MW: A6C0333D4CDEA03 CRC64;

Query Match 5.2%; Score 123.5; DB 1; Length 1858;
Best Local Similarity 27.5%; Pred. No. 0.21; Mismatches 38; Conservative 25; Indels 42; Gaps 5;
Matches 38;

DB 312 VNRDEKOFMH-LMNSFVRKQRYADGHISMACEAFSREYKELHRYSSLFMCWRLFILKL 370
DB 113 IEKKEKELINTLISNSGVTFYMT-----ALEILDSTINTPLNRSR----- 152
DB 371 WNGGLVDSATINNCNTILENCRS-----SDTTTNNNSVDRPSDSMTNNNTVDHP 423
DB 153 ---GSIGSMPI--CNMLTSSSSSSSTTATPTTSSNNNNNNNNNNNNNNNNNNNNNN 207
DB 424 NDIINKNNVNDKDNNSRD 441
DB 208 NNN 225

RESULT 6

S155_YEAST STANDARD: PRT: 1000 AA.

AC P43612; 01-NOV-1995 (Rel. 32, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE S154-associating protein SAP155.
GN SAP155 OR YFR040W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=96220458; PubMed=8649382;
RA Luke M.M., della Seta F., di Como C.J., Sugimoto H., Kobayashi R.,
RA Arndt K.T., "The SAP, a new family of proteins, associate and function positively
with the S154 phosphatase."
RL Mol. Cell. Biol. 16:2744-2755(1996).
RP SEQUENCE OF 98-1000 FROM N.A.
RC STRAIN=5288c / AB972;
RC MEDLINE=95400292; PubMed=7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shidrata T., Ozawa M.,
RA Sasakuma S.-I., Sasakuma M., Tsuchiya Y., Sueda E., Yokoyama K.,

```

OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE OF 1-843 FROM N.A.
RA Baladon V., Ballesta J.P.G., Bou G., del Rey F., Esteban P.F.,
RA Garcia-Cantalejo J.M., Garcia-Ramirez J.J., Gonzalez A., Jimenez A.,
RA Revuelta J.L., Santos M.A.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 737-1195 FROM N.A.
RA Gallion L., Dujon B.;
RL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
CC -1-SIMILARITY: STRONG, TO YEAST YII151C. ALSO TO THE N-TERMINAL OF
CC YEAST ALDEHYDE DEHYDROGENASE, MITOCHONRIAL 1 (ALD1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between The Swiss Institute of Bioinformatics and the EMBL outstation -
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z28321; CAA82176.1; -.
DR PIR; S38174; S38174.
DR SGD; S0001804; YKR096W.
KM Hypothetical protein; ATP-binding.
FT DOMAIN 210 250 ASN-RICH.
FT NP_BIND 465 472 ATP (POTENTIAL)..
SQ SEQUENCE 1195 AA; 137490 MW; D7251DEBD5523622 CRC64;

Query Match 5.0%; Score 120; DB 1; Length 1195;
Best Local Similarity 21.4%; Pred. No. 0.22;
Matches 62; Conservative 37; Mismatches 99; Indels 92; Gaps 10;

QY 166 RNNTRLRKVCFLPDSPSLNGTENG-I-TLLDGNRGIGYPEATELAGQ----- 213
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 42 RNTVANTLSLDFGYLVNPPSSKRQNSNSYVDINSKKRISRPRFSDIGKKNNDHTYPERTY 101

QY 214 -FEETSNIPRALHASSLDAGAKVLTISEAVYPARKTKKL-SAENSEARSHILLQOKROYH 271
   || | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 102 VKSEKMPSPRYVSSS-----KRLKRENSGVTGIQSALLISKSPENSGSJAHKR--WSP 154

QY 272 SHRPQPALEGVMSDDROSEDEVDDVDVAFEDRQLADLFVPDYVRKEQGFMLHANSFYRKOR 331
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 155 ENMKIKPLNVSO-----NSLA-----FVDAGSDPSQSKSELVGFGFKRSN 192

QY 332 VIADGHISMACEAFREFEKEELARYSSLFWCMWRLLFLIKLNHGVLVSATINNCNTIILENC 391
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 -----NSGEINDKD---NSA 204

QY 392 RNSSDDTTTNNNNSVDRPDSNTNNNNIVDHPDINNKNVYNKDKNNSRD 441
   | : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 RDQDFNNSGNHHNNHSSNNNDNNNNNDNNNDNNNNNSNRDN--NNSSDD 253

RESULT 8
YB00_YEAST STANDARD; PRT; 1094 AA.
AC P38114;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative 126.9 kDa transcriptional regulatory protein in YSW1-RIB7
DE Intergenic region.
GN YBR150C OR YBR1133.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
NX NCBI_TaxID=4932;
NN [1]
```

SEQUENCE FROM N.A.

RP STRAIN-5288C.

RA Ntjian K.-D., Koetter P., Rose M., Becker J., Grey M., Li Z.,

RA Nigemann E., Schenk-Groeninger R., Servos J., Wehner E.,

RA Wolter R., Brendel M., Bauer J., Braun H., Dorn K., Diestelhorst S.,

RA Grebenstein R., Hedges D., Kriesau P., Korol S., Krebs B., Proft M.,

RA Steiger K., Baur A., Boles E., Miosga T.,

RA Schmitt-Gerstenschlaeger I., Zimmermann F.K.;

RA Submitted (Aug-1994) to the EMBL/GenBank/DBJ databases.

CC -1- SIMILARITY LOCATION: Nuclear (Probable).

CC -1- SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINDING

CC CLUSTER DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its use by

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (see http://www.isb-sib.ch/announcement/or_send_an_email_to_license@isb-sib.ch).

CC -----

CC EMBL, Z36019; CA85108.1; -

CC PIR, S46021; S46021.

CC HSSP, P08657; 1CID.

CC TRANSFAC, T03677; -

CC SCD, S0000354; YBR150C.

CC InterPro, IPR001138; Fungal-Trn.

CC Pfam, PF00172; Zn-clus. 1.

CC SMART, SMO0066; GAL4. 1.

CC PROSITE, PS00463; ZN2_CYS6_FUNGAL.1; 1.

CC PROSITE, PS50048; ZN2_CYS6_FUNGAL.2; 1.

CC Hypothetical protein: Transcription regulation. DNA-binding;

CC Nuclear protein; Zinc: Metal-binding

CC DNA_BIND 107 137

CC SEQUENCE 1094 AA; 126902 MW; 3C0BF83F5879D14F CRC64;

QY Query Match 5.0%; Score 119; DB 1; Length 1094;

QY Best Local Similarity 27.5%; Pred. No. 0.23;

QY Matches 44; Conservative 26; Mismatches 44; Indels 46; Gaps 9;

QY 303 ROMIDDFVYVKKDEKQFHMNLNSFVRKQVIADGHISWACE-----AFSREYE--KE 352

QY 859 RESKDQWELKK-----TSINNEKMAFG--SFVTDQDKMEVSFIFNEITIQD 907

QY 353 LH--RYSLFQWCRFLFIKLINHGILVDSATINNCNTILLENCRNSDDTTTNNNS--VDRP 409

QY 908 LNFQFSS-----IPKLEWNTLEPG-----EEYHSHSGTGTDTNNETTGADDT 950

QY 410 SDSNTNNNNIYDHPNDI-----NNKNVNDKNNNSRDKVI 444

QY 951 DDNNNNNNNNNNKNGNNSSTINNNNNYSNNNDNDNNI 990

RESULT 9

OMB_DROME STANDARD; PRT; 988 AA.

AC 024432; Q27917; Q5W4K5;

AC 01-NOV-1997 (Rel. 35, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Opomator-blind protein (lethal(1)opomator-blind) (L(1)omb) (B1fid

DE protein).

GN BI OR OMB OR CG3578.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;

OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;

OX (1)

RP SEQUENCE FROM N.A.

RP STRAIN-Oregon R; TISSUE-Embryo;

RC MEDLINE=92159016; PubMed=1741374;

RA Pfingstfeldt G.O., Roth H., Poock B., Kerschner S., Schwarz H.,

RT Jongscheek B., Heisenberg M.: "The lethal(1)optomotor-blind gene of *Drosophila melanogaster* is a
RT major organizer of optic lobe development: Isolation and
RT characterization of the gene".
RL Proc. Natl. Acad. Sci. U.S.A. 69:1199-1203(1992).
RN [2].

PP SEQUENCE FROM N.A.
RC STRAIN-BECKER;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Chapple M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basay A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beszon K.Y., Berens P.V., Bernier B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Casway S., Dahlke C., Daveport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Necha S., Dunkov B.C., Dunn P.,
RA Foulger K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodde A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kaulush F., Karpen G.H., Ke Z., Kennison J.G., Ketchem R.A.,
RA Kimmel B.E., Kohtala C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laiko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McKernson D.,
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mosheiff A.,
RA Mount S.M., Moy M., Murphy B., Kurpyl L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Kusken D.R., Pacled J.M.,
RA Palazzolo M., Petlan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K.E., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sliden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spivker E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.Y., Wasserman D.A., Welleslock G.M., Weissensbach J.,
RA Ye J., Yen R.-F., Zaveril J.S., Zhao K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng Y.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Zheng L.,
RA Gladys R.A., Myers E.W., Rubin G.M., Venter J.C.;
RN "The genome sequence of *Drosophila melanogaster*".
RL Science 287:2185-2195(2000).

RN [3]
RP SEQUENCE OF 1-447 FROM N.A., AND MUTATIONAL ANALYSIS.
RQ TISSUE-LARVA;
RX MEDLINE=93261414; Pubmed=8492800;
RA Poock B., Balles J., Pfafffelder G.O.;
RT "Transcript identification in the optomotor-blind locus of *Drosophila*
RT melanogaster by intragenomic recombination mapping and PCR-aided
RT sequence analysis of lethal point mutations";
RL Mol. Gen. Genet. 238:325-332(1999).

CC -1- FUNCTION: ESSENTIAL PROTEIN THAT MAY FUNCTION AS A TRANSCRIPTION
CC REGULATOR. FLIES WITH LI(1)OMV MUTATIONS SHOW SEVERE MALDEVELOPMENT
CC OF THE OPTIC LOBES. REDUCTION IN KING SIZE AND AN INCREASED
CC ABDOMINAL PIGMENTATION. THEY DIE DURING THE PUPAL STAGE.

CC -1- SUBCELLULAR LOCATION: Nuclear (potential).

CC -1- TISSUE SPECIFICITY: IN THIRD INSTAR LARVAE IT IS FOUND IN THE
CC BRAIN REGION THAT WILL DEVELOP INTO OPTIC LOBES AND MORE WEAKLY IN
CC THE THORACIC PART OF THE VENTRAL GANGLION.

CC -1- DEVELOPMENTAL STAGE: THE PEAK PERIODS OF EXPRESSION ARE: MID-
CC EMBRYOGENESIS, THE SECOND DAY OF PUPAL DEVELOPMENT AND IN THE
CC ADULT.

CC -1- SIMILARITY: CONTAINS 1 T-BOX DOMAIN.

CC -----

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; M81796; AAA28736.1; -
 DR EMBL; AE003431; AAF45946.1; -
 DR EMBL; S61732; AAB26697.1; -
 DR EMBL; S61727; AAB26697.1; JOINED.
 DR EMBL; S61729; AAB26697.1; JOINED.
 DR EMBL; S61744; AAB26699.1; -
 DR EMBL; S61743; AAB26699.1; JOINED.
 DR EMBL; S61955; AAB26699.1; JOINED.
 DR HSSP; P24781; 1XBR.
 DR FLYBASE; FBgn0000179; b1.
 DR InterPro; IPR001699; TF_T-box.
 DR Pfam; PF00907; T-box; 1.
 DR PRINTS; PR00937; TBOX.
 DR SMART; SM00425; TBOX; 1.
 DR PROSITE; PS01283; TBOX_1; 1.
 DR PROSITE; PS01264; TBOX_2; 1.
 DR PROSITE; PS0252; TBOX_3; 1.
 KM DNA-binding; Nuclear protein; Transcription regulation.
 FT DOMAIN 51 140 ASN-RICH.
 FT DOMAIN 104 107 POLY-THR.
 FT DOMAIN 179 184 POLY-SER.
 FT DOMAIN 229 236 POLY-GLN.
 FT DOMAIN 238 244 POLY-PRO.
 FT DNA-BIND 332 513 T-BOX.
 FT DOMAIN 574 577 POLY-ASP.
 FT DOMAIN 607 692 ALA-RICH.
 FT DOMAIN 823 831 POLY-GLY.
 FT DOMAIN 910 916 POLY-ALA.
 FT DOMAIN 926 966 GLN/HIS-RICH.
 FT CONFLICT 10 10 F->L (IN REF. 1 AND 3).
 FT CONFLICT 216 216 A->P (IN REF. 1 AND 3).
 FT CONFLICT 511 511 F->L (IN REF. 1).
 FT CONFLICT 823 823 MISSING (IN REF. 1).
 FT CONFLICT 976 988 MISSING (IN REF. 1).
 SQ SEQUENCE 988 AA; 103992 MW; 032B7A4471743FC9 CMC64;

Query Match 4.9%; Score 117.5; DB 1; Length 988;
 Best Local Similarity 35.2%; Pred. No. 0.26;
 Matches 32; Conservative 11; Mismatches 23; Indels 25; Gaps 3;

QY 372 NHGLVDSATNNCNLENCN-----SSDTTNNNSVDSPD 411
 DB 59 NSGNNSNSNNNTNTNTNLVAVSPGGAQLSPQSSNHSNTTTTNTNTN-----SS 114
 QY 412 SNTNNNNIVDHPNDINNNKNDKDNNSRDK 442
 DB 115 SNNNNNNST-HNNNNNNHNNNNNNNNNNNTSOK 144

RESULT 10
 HMCU_DROME STANDARD; PRT; 2175 AA.
 AC P10180: Q9W306; (Rel. 10, Created)
 DT 01-MAR-1989 (Rel. 10, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Homeobox protein cut.
 GN CT OR CG11387.
 OS Drosophila melanogaster (Fruit fly).
 CC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 CC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 CC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88232956; PubMed=2897632;
 RA Blochlinger K., Bodmer R., Jack J., Jan L.Y., Jan Y.N.;
 RT "Primary structure and expression of a product from cut, a locus
 RT involved in specifying sensory organ identity in Drosophila.";

RL Nature 333:629-635(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikler S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blake R.G., CHANGE M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
 RA Abrell J.F., Agbayani A.I., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lal Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazono M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: REGULATOR OF CELL FATE DECISIONS IN MULTIPLE LINEAGES.
 CC SPECIFICALLY, FUNCTIONS AS A DETERMINATION FACTOR THAT SPECIFICS
 CC SENSOR ORGAN IDENTITY IN PRECURSOR CELLS. PROBABLY ALSO INVOLVED
 CC IN CELL TYPE SPECIFICATION OF MALPIGHIAN TUBULES. IN ABSENCE OF
 CC CUT GENE EXTERNAL SENSOR ORGANS ARE TRANSFORMED INTO CHORDOTONAL
 CC ORGANS.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: DETECTED IN MANY CELLS IN THE CENTRAL NERVOUS
 CC SYSTEM, ALL EXTERNAL SENSOR ORGANS, SOME PERIPHERAL NEURONS, AND
 CC IN THE NON-NEURAL CELLS OF THE SPIRACLES AND THE MALPIGHIAN
 CC TUBULES.
 CC -1- DEVELOPMENTAL STAGE: CELL-SPECIFIC PATTERN OF EXPRESSION. BROADLY
 CC EXPRESSED DURING EMBRYONIC DEVELOPMENT.
 CC -1- DOMAIN: ASN AT POSITION 47 OF THE HOMEBOX MAY PARTICIPATE IN
 CC REGULATING DNA-BINDING ACTIVITY BY PROMOTING HOMO- AND
 CC HETERO-DIMERIZATION.
 CC -1- SIMILARITY: BELONGS TO THE CUT HOMEBOX FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

[illegible]

Query Match	4.98;	Score 117;	DB 1;	Length 2175;
Best Local Similarity	23.18;	Pred. No. 0.8;		
Matches 58;	Conservative 32;	Mismatches 107;	Indels 54;	Gaps 8

```

QY 220 I P P A I H S S I D A G A K V I L T S E A V V A T K R L S A E R S E A R S H L L O K Q O F Y S H R V O P M A 279
    |||
Dh 222 I P T R O P A Y A T A T A A A A A A A A T P I A T O N N A S G S T T S N A N ----- H T S N N S H Q 270
    |||
QY 280 L E O Y A S D R S D E D E V D D V A D F E D R - Q M L D F V Y N K E R Q O F M H L N N S F V R K O R Y A I G H I 338
    |||
Dh 271 D E E I D D E E D E E D E D E E E E N A S M O S N A D D E D I D A O E T R T E S A T T O O P ----- H Q 335
    |||
QY 339 S W A C E A F S R E Y E K E L H R Y S L P W C R L F L I K M H G I G V S A T I N N C A T I L N C S N S D - 396
    |||
Dh 326 Q O O D T D L E E - N K D A G E A S ----- L N S N N H N T T D S N - N S C S R K N N N G C N E S O H 372
    |||
QY 397 ----- T T T T N N N S V D R P D S S N T N N N N I Y D H E N D I N K N N Y N D K D N N ----- 438
    |||
Dh 373 V A S S A E D D C C A N N N T Y T S N N N T S T A T S N T N N N N N N N S S C G S E K E R K K K N N N N N G O P A 432
    |||
QY 439 ----- S R D K V I R 445
    |||
Dh 433 V I L A A K D R E I K 443
    |||

```

RESULT 11	Y1NO_YEAST	STANDARD:	PRT;	964 AA.
ID	Y1NO_YEAST			
AC	P40467;			
DT	01-FEB-1995 (Rel. 31, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Interative 108.8 kDa transcriptional regulatory protein in FKH1-STH1			
DE	negative region.			
GN	Y1H130W.			
OS	Saccharomyces cerevisiae (Baker's yeast).			
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;			
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyc.			

```

OX  NBI_TaxID=4932;
X  [1]
RP  SEQUENCE FROM N.A
RC  STRAIN=5286 / AB972;
RA  Bartell B.G., Badcock K., Bankier A.T., Bowman S., Brown D.,
RA  Churchill C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
RA  Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
RA  Louis E., Lyne G., Moule S., Moule T., Odell C., Pearson D.,
RA  Rajadream M.A., Riles L., Rowley N., Skelton J., Smith V.,
RA  Walsh S.V., Whitehead S.;
RL  Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
CC  -1-SUBCELLULAR LOCATION: Nuclear (Probable) ..
CC  -1-SIMILARITY: CONTAINS 1 ZN(2)-CYS(6), FUNGAL-TYPE BINUCLEAR
CC  CLUSTER DOMAIN.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation at
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/)
CC  or send an email to license@isb-sib.ch.

```

CC EMBL; Z38059; CAA86148.1; -.
DR PIR; S48404; S48404.
DR HSSP; P08657; 1CID.
DR SGD; S0001392; YII130W.
DR InterPro; IPR001138; FunG1_T1N.
DR Pfam; PF001172; Zn_C1us; 1.
DR PRINTS; PRO0054; FUNGALZNCYS.
DR SMART; SM00066; GAL4; 1.
DR PROSITE; PS00463; ZN2_CY6_FUNGAL_1; 1.
DR PROSITE; PS00468; ZN2_CY6_FUNGAL_2; 1.
KW Hypothetical protein; Transcription regulation; DNA-binding;
KW Nuclear protein; Zinc; Metal-binding.
FT DNASIS_BIND 21 47 ZN(2)-CYS(6), FUNGAL-TYPE.
FT DOMAIN 811 896 ASN-RICH.
SQ SEQUENCE 964 AA; 108780 MW; AD5ABE59E4B022CC CRC64;

	Query Match	4.8%;	Score 114;	Ds1:	Length 964;	
	Best Local Similarity	40.4%;	Pred. No. 0.47;			
	Matches 23; Conservative	11;	Mismatches	21;	Gaps	1;
Qy	382	NNCNTLLERCRSSDSTTTNNNSVYDRSDSTENNNTTVDHPADINKKNVYDKDN	438			
		: :				
Dz	833	NNNNNTTNNSNNNNNDNDDNNNNNNNNNNNNNGGSSN--NNNNNNNNKKNN	887			

	FAB1_YEAST	STANDARD:	PRT; 2278 AA.
FP	FAB1_YEAST		
IC	P34756:		
DT	01-FEB-1994 (rel. 28, Created)		
DT	01-FEB-1994 (rel. 28, Last sequence update)		
DT	15-JUN-2002 (rel. 41, Last annotation update)		
DE	Probable phosphatidylinositol-4-phosphate 5-kinase FAB1 (EC 2.7.1.66)		
DE	(1-phosphatidylinositol-4-phosphate kinase) (PIP5K) (PtdIns(4)P-5-		
DE	kinase) (Dlpihosmositide kinase).		
GN	FAB1 OR YFR019W.		
OS	Saccharomyces cerevisiae (Baker's yeast).		
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
OC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.		
OX	NCBI_TaxID=4932;		
RP	[1]		
RP	SEQUENCE FROM N.A.		
RA	MEDLINE=95392039; PubMed=7663021;		
RA	Yamamoto A., Dewald D.B., Boronienkov I.V., Anderson R.A., Emr S.D.,		
RA	Koshland D.;		
RT	"Novel PI(4)P 5-kinase homologue, Fabip, essential for normal vacuole		
RT	function and morphology in yeast.";		
RL	Mol. Biol. Cell 6:525-539(1995).		
RP	[2]		
RP	SEQUENCE FROM N.A.		


```

RC STAIN-S288C / AB972;
RX MEDLINE-95400292; PubMed-7670463;
RA Murakami Y., Naitou M., Hagiwara H., Shibata T., Ozawa M.,
RA Sasanuma S.-I., Sasanuma M., Tsuchiya Y., Soeda E., Yokoyama K.,
RA Yamazaki M., Tashiro H., Eki T.;
RT "Analysis of the nucleotide sequence of chromosome VI from
RT Saccharomyces cerevisiae ";
RL Nat. Genet. 10:261-268(1995).
RM [3]
RN SIMILARITY TO TCP1/CCP FAMILY.
RP MEDLINE-95314774; PubMed-7794526;
RX Waldmann T., Lupas A., Kellermann J., Peters J., Baumeister W.;
RA "Primary structure of the thermosome from Thermoplasma acidophilum.";
RL Biol. Chem. Hoppe-Seyler 376:119-126(1995).
CC -1- FUNCTION: CATALYZES THE PHOSPHORYLATION OF PHOSPHATIDYLINOSITOL-4-
CC PHOSPHATE ON THE FIFTH HYDROXYL OF THE MYO-INOSITOL RING, TO FORM
CC PHOSPHATIDYLINOSITOL-4,5-BIPHOSPHATE. REQUIRED FOR ENDOCYTIC-
CC VACUOLAR PATHWAY AND NUCLEAR MIGRATION. THE PRODUCT OF THE
CC REACTION IT CATALYZES FUNCTIONS AS AN IMPORTANT REGULATOR OF
CC VACUOLE HOMEOSTASIS PERHAPS BY CONTROLLING MEMBRANE FLUX TO AND/OR
CC FROM THE VACUOLE.
CC -1- CATALYTIC ACTIVITY: ATP + 1-phosphatidy1-ID-myo-inositol 4-
CC monophosphate = ADP + 1-phosphatidy1-ID-myo-inositol 4,5-
CC biphosphate.
CC -1- SUBCELLULAR LOCATION: VACUOLE-ASSOCIATED (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE PTINIS(4)P-5-KINASE FAMILY.
CC -1- SIMILARITY: CONTAINS 1 FYVE-TYPE ZINC FINGER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U01017; AAA81360.1; -
DR EMBL: D50617; BAA09258.1; -
DR SGD: S0001915; FABI.
DR InterPro: IPR002498; PIP5K.
DR InterPro: IPR000306; Znf_FYVE.
DR Pfam: PF01363; FYVE.1.
DR SMART: SM0064; FYVE; 1.
DR SMART: SM00330; PIPKC; 1.
DR PROSITE: PS0178; ZF_FYVE; 1.
DR Trifinger: PS0178; ZF_FYVE; 1.
FT ZN_FING 240 299 FYVE-TYPE.
FT DOMAIN 393 397 POLY-PRO.
FT DOMAIN 571 590 POLY-ASN.
FT DOMAIN 1808 1811 POLY-THR.
FT DOMAIN 1891 1897 POLY-GLN.
FT CONFLICT 2275 2275 R -> W (IN REF. 2).
FT SEQUENCE 2278 AA; 257417 MW; 1A0A30E13165DE41 CRC64;

```

[illegible]

CC -1 SUBUNIT: COMPONENT OF THE SWI/SNF GLOBAL TRANSCRIPTION ACTIVATOR
 CC COMPLEX.
 CC -1 SUBCELLULAR LOCATION: Nuclear.
 CC -1 SIMILARITY: CONTAINS 1 ARID DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U66528; AAB06791.1;
 CC Dictydb: DD01058; c19b.
 CC InterPro: IPR000315; Znf_Box.
 CC Pfam: PF00643; zif-B_box; 1.
 CC Repeat.
 KW

Query Match
 Best Local Similarity 4.7%; Score 112; DB 1; Length 1314;
 Matches 33; Conservative 11; Mismatches 28; Indels 22; Gaps 5;
 Db 372 NGLVDSATINNCNTILENCRNSSD-TTNNNSVDRPS--DSNTNNN-
 Db 13 NNTTTTNTNNNNNTNNNNNTNNNNPANNNTNNNSSTNNNTNNNTNTGASCVD 72
 Qy 422 HNN-----DNNKN-NVDKNDSRDKY 443
 Db 73 FQNFDPKPFQNDLDSNNNSNNNSNNNNNNNTY 106

RESULT 14
 CIGB_DICDI STANDARD; PRT; 735 AA.
 AC Q94481;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein c19b (Fragment).
 GN CIGB.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Loomis W.F.;
 RC STRAIN=AX4;
 RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
 CC -1 SIMILARITY: TO D.DISCOIDEUM CIGB.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U66528; AAB06791.1;
 CC Dictydb: DD01058; c19b.
 CC InterPro: IPR000315; Znf_Box.
 CC Pfam: PF00643; zif-B_box; 1.
 CC Repeat.
 KW

FT NON TER 1 155 POLY-ASN.
 FT DOMAIN 151 251 POLY-ASN.
 FT DOMAIN 251 268 TANDEN REPEATS.
 SO SEQUENCE 735 AA; 83443 MW; 9765013088DA6C8C CRC64;
 Query Match
 Best Local Similarity 19.9%; Score 111.5; DB 1; Length 735;
 Matches 73; Conservative 56; Mismatches 117; Indels 121; Gaps 17;
 Db 87 DCSGPFSCMGSGFKGLPHLNSHDLFEFEKLEFEQTVNVSXKNSPFEFGSDDP 146
 Db 20 DCKFMCCSLCTSRKG-EHHGKTDLS-----ESSASNIHSLNDF-----KD 61
 Qy 147 KPEPFLSCPRKRRGGRNTRKLVKCFPLDPSLTNGTNGTILNDGNGIGYPE 206
 Db 62 NIYP-----KLEKKE-----NDQILKE-----SNETPKELIQSDNDNNL 98
 Qy 207 ATEIAGQPEMTNIPPAI-----AHSIDGAKV--ILTSAAVPAKTKLSAENSE 257
 Db 99 ---LNKEFKQIHNLSEVLEIDIEKQISTNDONILNITTS-----INNDKIIITLN 151
 Qy 258 ARSHLLCKROPYSHRYQPMALEROVSDSEDEVDVADVEDROMIDPVYNNKDEK 317
 Db 152 NNNNSIIDLKLYF-----TQINNOODNDNDQCYQIDPTIELIKQY-----QN 195
 Qy 318 QFMHLMNSFYRKQVINDGHSVACAFSFRFEKELHRYSSLFCMCRFLIKLNHGLV 376
 Db 196 SLITIKNS-----NNIN-----NLTKLENNQILK 221
 Qy 377 -DSATINNCNTILENCRNSSDPTTNNNSVDRPSDNTNNNNIYDHPNDNNKNVNDK 435
 Db 222 FDMQNTINIKNSFKSIYSQDLP-----NDALDYNNNNNNNN-----NNNNNNNE 269
 Qy 436 DNNSRDK 442
 Db 270 LNNNNK 276

RESULT 15
 CYAG_DICDI STANDARD; PRT; 858 AA.
 AC Q03101;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adenylate cyclase, germination specific (Ec 4.6.1.1) (ATP
 DE pyrophosphate-lyase) (Adenyl1 cyclase).
 GN ACQA OR AGC.
 OS Dictyostelium discoideum (Slime mold).
 OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 OX NCBI_TaxID=44689;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92233467; PubMed=1348970;
 RA Platt G.S., Milona N., Borleis J., Lin K.C., Reed R.R.,
 RA "Structurally distinct and stage-specific adenyl1 cyclase genes play
 RA different roles in Dictyostelium development.";
 RL Cell 69:305-315(1992).
 RN [2]
 RP 3D-STRUCTURE MODELING OF 387-543.
 RA MEDLINE=98054247; PubMed=9391039;
 RA Liu Y., Runbo A.B., Rao V.D., Hurley J.H.;
 RA "Catalytic mechanism of the adenyl1 and guanylyl cyclases: modeling
 RA and mutational analysis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 94:13414-13419(1997).
 CC -1 FUNCTION: CG HAS A LARGE EXTRACELLULAR DOMAIN WHICH MAY BE
 CC INVOLVED IN THE RECOGNITION OF AN EXTRACELLULAR SIGNAL PRESENT
 CC DURING GERMINATION, LEADING TO ACTIVATION OR INHIBITION OF CAMP
 CC CATALYTIC ACTIVITY: ATP -> 3',5'-cyclic AMP + diphosphate.
 CC -1 ENZYME REGULATION: INSENSITIVE TO GUANINE NUCLEOTIDES.
 CC

```
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.
CC -1- DEVELOPMENTAL STAGE: AFTER FRUITING BODIES HAVE BEEN FORMED AND
CC DURING GERMINATION.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 CHASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M87278; AAA33164.1; -
DR PDB: 1AMW; 28-JAN-98.
DR DictyDB; DD04007; acga.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF00211; guanylate_cyc; 1.
DR SMART; SM00044; CYCC; 1.
DR PROSITE; PS50839; CHASE; 1.
DR PROSITE; PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE; PS50125; GUANYLATE_CYCLASES_2; 1.
DR Lyase; CAMP synthesis; Transmembrane; Germination; signal-anchor;
KW 3D-structure.
FT DOMAIN 1 18 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 19 41 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 42 858 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 86 317 CHASE.
FT DOMAIN 396 526 GUANYLATE CYCLASE.
FT DOMAIN 738 848 ASN-RICH.
FT DOMAIN 835 848 POLY-ASN.
SQ SEQUENCE 858 AA; 98413 MW; B483FB255289E65D CRC64;

Query Match 4.7%; Score 111.5; DB 1; Length 858;
Best Local Similarity 19.2%; Pred. No. 0.64;
Matches 76; Conservative 57; Mismatches 135; Indels 127; Gaps 15;

QY 104 QPHLNS-HDLFEPEFKLEFEYQIVNSVK--LNSFEEEGSDDKFEPPSLCSKPRK 159
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 533 QIHVSDRYQLGKEDFNFSERCDIHVKGGMKTWYLMGKSSDPSLKKDFS-----RS 587
QY 160 RRORGRRNTRRLKCYFLPLDSPSLTNGTENGITLLDGNRGLGYPEATELAGOFEMTSN 219
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 588 RVO-----PSLFNRKQSHVCI-----YPEPPSGIQLALNIENN 620
QY 220 IPPAIHSSLDAG---AKVILTSEAVVPATKTRKLSAERSPARSHLLQKROFYHSHRV 275
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 621 L-----NNTDAGCENCCKILKTYAYSPDST-----SNYYTHGDN 657
QY 276 QPMALQVMSDR-DSEDEVDDVADPEDROMLIDFVYVKDEKQPMH--LMSFVYRKQRY 332
   :::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 658 SPPEPSLNSNDLIDSEYHDHPFSPDSNVGYHDTSKDIKEDENQNETLLFNOQLKKQ 717
QY 333 IADGHSWACEAFSFFYEKELHRYSLFWCWRLLFLIKLMNHGLVDSATLNNCNTILLENCR 392
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 718 IEN-----IQRDLSDLS-----IEAKILNNNNNNNNINDNNNTNPNNN 758
QY 393 -----NSSDTT---NNNNSVDRPSDSNTNNNTIVD----- 422
   |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 759 NNDIINNSDNNYNNNNNSDKLENNDDGNNNINDNNKYSTNENNIKSKTLFQDSKL 818
QY 423 -----PNDINKKNVNDKNNRSRDKVYK 445
   |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
DB 819 INDIKAKENCNDNDNNNNNNNNNNNNNNNDENVE 853
```

Search completed: June 13, 2003, 15:36:11
Job time : 25 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:45:40 ; Search time 4487 Seconds

(without alignments)
11168.939 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctcttcaattgtct.....aatgattctgtcttaact 1722

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_ph:*

7: gb_pl:*

8: gb_pr:*

9: gb_ro:*

10: gb_ro:*

11: gb_sy:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1722	100.0	1722	6	AX032817	AX032817 Sequence
2	1722	100.0	1722	6	AX032888	AX032888 Sequence
3	1720.4	99.9	1722	6	AX032934	AX032934 Sequence
4	1678.6	97.5	1700	8	AF284500	AF284500 Arabidops
5	1646.6	95.6	1722	8	AF284501	AF284501 Arabidops
6	1569.6	91.1	1715	6	AX032818	AX032818 Sequence
7	1569.6	91.1	1715	6	AX032891	AX032891 Sequence
8	1568.6	91.1	1721	8	AY034902	AY034902 Arabidops
9	1537.6	89.3	1354	6	AY063047	AY063047 Arabidops
10	1277.4	74.2	1354	8	AX032894	AX032894 Sequence
11	412.2	23.9	6338	6	AX032890	AX032890 Sequence
12	407.8	23.7	91660	8	ATRP51E2	AF180942 Arabidops
13	319.4	18.5	5895	6	AX032893	AX032893 Sequence
14	319.4	18.5	199348	8	ATCRIV45	AL161545 Arabidops
15	319.4	18.5	201471	8	ATRCAT7	Z97342 Arabidops
16	291.4	16.9	2279	8	AB053171	AB053171 Arabidops
17	236.2	13.7	2384	8	AF096096	AF096096 Arabidops
18	230.8	13.4	558	6	AX032900	AX032900 Sequence
19	163.8	9.5	325	6	AX032898	AX032898 Sequence
20	126.6	7.4	463	6	AX032896	AX032896 Sequence
21	78.4	4.6	35276	2	AC115611	AC115611 Dictyoste
22	75.2	4.4	256774	2	AC116964	AC116964 Dictyoste
23	73.4	4.3	106434	3	AC117080	AC117080 Dictyoste
24	73.4	4.3	147192	2	AC116925	AC116925 Dictyoste
25	73.2	4.3	79554	8	NCB11B23	AL669991 Neurospor
26	72.8	4.2	141008	2	AC126897	AC126897 Rattus no
27	72.8	4.2	144022	2	AC109949	AC109949 Rattus no
28	72.8	4.2	185424	2	AC096166	AC096166 Rattus no
29	72.2	4.2	166918	2	AC128567	AC128567 Rattus no
30	72.2	4.2	237563	3	AF153362	AF153362 Dictyoste
31	71.8	4.2	79291	2	AF153362	AF153362 Dictyoste
32	71.8	4.2	97683	2	AC116348	AC116348 Dictyoste
33	71.8	4.2	212383	2	AC119030	AC119030 Rattus no
34	71.6	4.2	57906	8	NCB1J19	AL389899 Neurospor
35	71.6	4.1	656	3	AF453312	AF453312 Cotesia c
36	71.4	4.1	8045	6	AX034425	AX034425 Sequence
37	71.4	4.1	156533	2	AC117070	AC117070 Dictyoste
38	71.4	4.1	229502	2	AC126871	AC126871 Rattus no
39	70.6	4.1	153359	2	AC096792	AC096792 Rattus no
40	70.6	4.1	165337	2	AC096212	AC096212 Rattus no
41	70.6	4.1	176351	2	AC117007	AC117007 Rattus no
42	70.6	4.1	179380	2	AC120325	AC120325 Rattus no
43	70.4	4.1	2790	3	DD301670	AJ301670 Dictyoste
44	70.4	4.1	6016	8	AY077730	AY077730 Podospora
45	70.4	4.1	53440	2	AC117078	AC117078 Dictyoste

ALIGNMENTS

RESULT 1

AX032817 1722 bp DNA 1linear PAT 21-SEP-2000

LOCUS AX032817

DEFINITION Sequence 57 from Patent WO0046558.

ACCESSION AX032817

VERSION AX032817.1 GI:10279793

KEYWORDS

SOURCE

ORGANISM

synthetic construct.

synthetic construct

artificial sequences.

REFERENCE

1 (bases 1 to 1722)

AUTHORS Dean,C., West,J. and Johanson,U.

TITLE Plant gene

JOURNAL Patent: WO 0046358-A 57 10-AUG-2000;

DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;

Pred. No. is the number of results predicted by chance to have a

Fri Jun 20 09:03:38 2003

us-09-890-220-1.rge

Page 2

JOHANSON	URBAN (SE)
FEATURES	Location/Qualifiers
source	1. .1722

BASE COUNT	ORIGIN
538 a	339 c 348 g 497 t

Query Match	Score	DB 6	Length
100.0%	1722	1722	1722

```
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY	1	CAAGCTCTTCATTTTGGTTCCTCTCTCTTACACAGCCAAATCGGTGTTTCCGACCTTT	60
Db	1	CAAGCTCTTCATTTTGGTTCCTCTCTCTTACACAGCCAAATCGGTGTTTCCGACCTTT	60
QY	61	CAGGCTCAATCCAAAGACATTCATATATTAAGCATATTTGAGAAGAGCGGTTCTAATTTGT	120
Db	61	CAGGCTCAATCCAAAGACATTCATATATTAAGCATATTCAGAAAGGCGGTTCTAATTTGT	120
QY	121	GCATTAAGTTATTCGGTATACGATGAGGGAATTTCAATTTAGGGAGAGCCTCAGATTTG	180
Db	121	GCATTAAGTTATTCGGTATACGATGAGGGAATTTCAATTTAGGGAGAGCCTCAGATTTG	180
QY	181	CATTAACCTCATATATGGGCTCTTGACGTTGTTGAGTGTAATTTGAACAAAGATGTGAGGC	240
Db	181	CATTAACCTCATATATGGGCTCTTGACGTTGTTGAGTGTAATTTGAACAAAGATGTGAGGC	240
QY	241	AGAAATGTGCGCGGAATCTCTACCGGAGAAATGATTTCAACTGATGAGATCTCTTGA	300
Db	241	AGAAATGTGCGCGGAATCTCTACCGGAGAAATGATTTCAACTGATGAGATCTCTTGA	300
QY	301	TATATTTGTAACCTGTGGACATATATTAACATCTTTCACCTTCGCTCTAGGCAACCAT	360
Db	301	TATATTTGTAACCTGTGGACATATATTAACATCTTTCACCTTCGCTCTAGGCAACCAT	360
QY	361	CGTTTCTTCCAAAGATGCTTGAACATAATTTGAGCAAAACCGCAAAAGAAATCAAGAT	420
Db	361	CGTTTCTTCCAAAGATGCTTGAACATAATTTGAGCAAAACCGCAAAAGAAATCAAGAT	420
QY	421	CTATCTGGGANGGATGTTTCAACTATTAAGATTTGATTAATACACTTTCACAAAATCTGAAG	480
Db	421	CTATCTGGGANGGATGTTTCAACTATTAAGATTTGATTAATACACTTTCACAAAATCTGAAG	480
QY	481	TTTAGGAGAGATTTGTTCTTGTCATTTGCTCTATGCTATGTGTAATGTAAGGAGCTGC	540
Db	481	TTTAGGAGAGATTTGTTCTTGTCATTTGCTCTATGCTATGTGTAATGTAAGGAGCTGC	540
QY	541	AATTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAACCTTTTGAGAAAT	600
Db	541	AATTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAACCTTTTGAGAAAT	600
QY	601	ACGAGACATTAATGTTTCTGTAAAACCTAATTTCCCTCATATTTGAGAGAAAGAAAGTg	660
Db	601	ACGAGACATTAATGTTTCTGTAAAACCTAATTTCCCTCATATTTGAGAGAAAGAAAGTg	660
QY	661	ATGACATTAATTTGAGCCCTTCTCTCTGCTCGAAACCTCGTAACGGAGACAAGAG	720
Db	661	ATGACATTAATTTGAGCCCTTCTCTCTGCTCGAAACCTCGTAACGGAGACAAGAG	720
QY	721	GTGGCAGAATTAACACGAGAGACTTAAAGATGCTTTTACGTTGAGATTCACCCAGTT	780
Db	721	GTGGCAGAATTAACACGAGAGACTTAAAGATGCTTTTACGTTGAGATTCACCCAGTT	780
QY	781	TAACTTAATGGCAGAAAAATGGAATCACCCACTTAATGATGGAACCGAGGTTTAAAGAT	840
Db	781	TAACTTAATGGCAGAAAAATGGAATCACCCACTTAATGATGGAACCGAGGTTTAAAGAT	840
QY	841	ATCCGAGGCAACAGAGCTTGCTGACATTTGAGATGACACGACATTTCCACAGCCA	900
Db	841	ATCCGAGGCAACAGAGCTTGCTGACATTTGAGATGACACGACATTTCCACAGCCA	900

[illegible]

the vernalization gene *Vrn2*
 Patent: WO 0044918-A 47 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)

FEATURES
 source Location/Qualifiers
 1..1722 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"

BASE COUNT 539 a 339 c 347 g 497 t

Query Match 99.98%; Score 1720.4; DB 6; Length 1722;
 Best Local Similarity 99.98%; Pred. No. 0;
 Matches 1721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CAAGCTTCTCAATTTGGCTGCTCTCTCTTACACAGCCATCGGTGTTGCGAGCTT 60
 Db 1 CAAGCTTCTCAATTTGGCTGCTCTCTCTTACACAGCCATCGGTGTTGCGAGCTT 60
 QY 61 CAGGCTCAATCCAGACATCTCTATATAGCATATGCGAAGAGGGGCTTCTATGTT 120
 Db 61 CAGGCTCAATCCAGACATCTCTATATAGCATATGCGAAGAGGGGCTTCTATGTT 120
 QY 121 GGATGAGTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 Db 121 GGATGAGTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 QY 121 GCATGAGTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 Db 121 GCATGAGTTATGCTATGAGTGAAGAAATTTCAATTTAGGGAGGCTCAGAGTTG 180
 QY 181 CACTAATCTCAATATGCGCTTGTGAGTGTGATTAATTGAACAAGAAATGTAGGC 240
 Db 181 CACTAATCTCAATATGCGCTTGTGAGTGTGATTAATTGAACAAGAAATGTAGGC 240
 QY 241 AGAATGTGCGCGAAATCTCCACCGGAGGAGAGTATTTCAACGATGAGAAATCTTGA 300
 Db 241 AGAATGTGCGCGAAATCTCCACCGGAGGAGAGTATTTCAACGATGAGAAATCTTGA 300
 QY 301 TATATGTAAACCTGTGAGATATATTAACATCTTTCACCTCGCTCTAGGCAACCAT 360
 Db 301 TATATGTAAACCTGTGAGATATATTAACATCTTTCACCTCGCTCTAGGCAACCAT 360
 QY 361 CGTTCTCCAGATAGCTGAACTCAAAATTTGGAGCAAGGCAAAAAGTCAAGAT 420
 Db 361 CGTTCTCCAGATAGCTGAACTCAAAATTTGGAGCAAGGCAAAAAGTCAAGAT 420
 QY 421 CTACTGGGATGATGTTTCAACATTAAGGATGTAATTAACACATTTACAGAAATCTAAG 480
 Db 421 CTACTGGGATGATGTTTCAACATTAAGGATGTAATTAACACATTTACAGAAATCTAAG 480
 QY 481 TTAGGAGGATGTTCTGTCCATTTTGTCTATGCTATGCTATGCTTCAAGGGCTGC 540
 Db 481 TTAGGAGGATGTTCTGTCCATTTTGTCTATGCTATGCTATGCTTCAAGGGCTGC 540
 QY 541 AATTCATTTGAATTCATCTCATGATTAATTGAATTTGAATTTGAATTTGAATTTGAAT 600
 Db 541 AATTCATTTGAATTCATCTCATGATTAATTGAATTTGAATTTGAATTTGAATTTGAAT 600
 QY 601 ACCAGACAGTAAATTTCTGTAAACCTAATTTCTCATTTGAGAGAAAGAGAGT 660
 Db 601 ACCAGACAGTAAATTTCTGTAAACCTAATTTCTCATTTGAGAGAAAGAGAGT 660
 QY 661 ATGAGCATTAATTTGAGCCCTTCTCTCTGCTGAAACCTCGTAAGCGGAGACAAAG 720
 Db 661 ATGAGCATTAATTTGAGCCCTTCTCTCTGCTGAAACCTCGTAAGCGGAGACAAAG 720
 QY 721 GTGGCAGAAATTAACACGAGAGACTTAAGATAGCTTTTACCGTTGATTCACCCAGTT 780
 Db 721 GTGGCAGAAATTAACACGAGAGACTTAAGATAGCTTTTACCGTTGATTCACCCAGTT 780
 QY 781 TAACATAATGACAGAAATGGAATCACCCTACTTAATGATGAGAAACCGGTGTTAGAT 840
 Db 781 TAACATAATGACAGAAATGGAATCACCCTACTTAATGATGAGAAACCGGTGTTAGAT 840
 QY 841 ATCCGAGGCAACAGGCTGTGTGCAATTTGATGACAGAAATTCACCAACCAACA 900
 Db 841 ATCCGAGGCAACAGGCTGTGTGCAATTTGATGACAGAAATTCACCAACCAACA 900

Db 841 ATCCGAGGCAACAGGCTGTGTGCAATTTGATGACAGAAATTCACCAACCAACA 900
 QY 901 TAGCCCATCTTCTCTGAGAGCTGCTCTAAATTAATTAATGACAGAAAGCTGCTGC 960
 Db 901 TAGCCCATCTTCTCTGAGAGCTGCTCTAAATTAATTAATGACAGAAAGCTGCTGC 960
 QY 961 CTGCTACTAGACAGAAAGTATCTCTGAGAGCTCAGAGCTAGAACCCACTTACTTC 1020
 Db 961 CTGCTACTAGACAGAAAGTATCTCTGAGAGCTCAGAGCTAGAACCCACTTACTTC 1020
 QY 1021 TTCAAGAAAGCCAAATTCATCATCTCAGAGCTCAGCCAAATGCGCTTGAGCAATGA 1080
 Db 1021 TTCAAGAAAGCCAAATTCATCATCTCAGAGCTCAGCCAAATGCGCTTGAGCAATGA 1080
 QY 1081 TGTCTGACCGGAGATAGGAGATGATGATGATGAATGAAGAAAGCAATTCATGATTTGA 1140
 Db 1081 TGTCTGACCGGAGATAGGAGATGATGATGATGAATGAAGAAAGCAATTCATGATTTGA 1140
 QY 1141 AGATGCTGATGATGATGATGATGATGATGAATGAAGAAAGCAATTCATGATTTGA 1200
 Db 1141 AGATGCTGATGATGATGATGATGATGATGAATGAAGAAAGCAATTCATGATTTGA 1200
 QY 1201 ACTGCTTTGTAAGAAACAAAGGTTTATGACAGATGCTATATCTCTGGCATGTGAAG 1260
 Db 1201 ACTGCTTTGTAAGAAACAAAGGTTTATGACAGATGCTATATCTCTGGCATGTGAAG 1260
 QY 1261 CATTTTCAAGATTTTACGAGAAAGGTTGACCGCTTACTCATCATCTCTCTGTGTGGA 1320
 Db 1261 CATTTTCAAGATTTTACGAGAAAGGTTGACCGCTTACTCATCATCTCTCTGTGTGGA 1320
 QY 1321 GATTTGTTTGAATTAATGATGAAACCATGACTTGTGACCTCAGCCACCATCAACACT 1380
 Db 1321 GATTTGTTTGAATTAATGATGAAACCATGACTTGTGACCTCAGCCACCATCAACACT 1380
 QY 1381 GCAATATCATCTCTGAGAGATGCGCTTAATGAGTGTGACACACACACCAACCAACA 1440
 Db 1381 GCAATATCATCTCTGAGAGATGCGCTTAATGAGTGTGACACACACACCAACCAACA 1440
 QY 1441 ACAGTGTGATGCTGCTGAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1500
 Db 1441 ACAGTGTGATGCTGCTGAGAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1500
 QY 1501 ATGACATTAACACAGAAAGCAATGTTGACAAACAGCAATTAACAGCAAGCAAGTAA 1560
 Db 1501 ATGACATTAACACAGAAAGCAATGTTGACAAACAGCAATTAACAGCAAGCAAGTAA 1560
 QY 1561 TTAATTAAGAAATCTCGGCTTTTATGATACGATTTATCGATTTGTAATTTCTTC 1620
 Db 1561 TTAATTAAGAAATCTCGGCTTTTATGATACGATTTATCGATTTGTAATTTCTTC 1620
 QY 1621 TTTCTTAAATAATGTTAGAGCAACAATTTTATATGTTAGTATCAACTGAT 1680
 Db 1621 TTTCTTAAATAATGTTAGAGCAACAATTTTATATGTTAGTATCAACTGAT 1680
 QY 1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGAT 1722
 Db 1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGAT 1722

RESULT 4
 AF284500 1700 bp mRNA linear PLN 21-NOV-2001
 LOCUS AF284500
 DEFINITION Arabidopsis thaliana vernalization 2 Protein (VRN2) mRNA, complete
 cds.
 ACCESSION AF284500
 VERSION AF284500.1 GI:16945787
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1700)

AUTHORS Gendall, A.R., Levy, Y.Y., Wilson, A. and Dean, C.
TITLE The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis
JOURNAL Cell 107 (4), 525-535 (2001)
MEDLINE 21575875
PUBMED 11719192
REFERENCE 2 (bases 1 to 1700)
AUTHORS Gendall, A.R., Levy, Y.Y. and Dean, C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
FEATURES
source
 1..1700
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="IV"
 /map="adjacent to RPP5"
 /note="ecotype: Landsberg erecta"
 1..1700
 /gene="VRN2"
 227..1564
 /gene="VRN2"
 /codon_start=1
 /product="vernalization 2 protein"
 /protein_id="AAL32135.1"
 /db_xref="GI:1694578"
 /translation="MCRONRAKSSPEEYISDENLITYCKPRLXNTPHLISGNPS
 FLPRCLNTYKIKRRKRKRSSTGVYFNKDCNNTLOKTEVEDSCPPCSMLGSGK
 LQFHLNSHDLFEFEKLESEYQTVNVSXKLSFLEEKSDDDFEPSLSCKPKR
 RQGRGNNTRLKVCFLPLDSPITNGENGLTLNDGNRGLGYEPAELAGOFKETS
 NIPPAIASHSLDAGAKVILTSAAVVPATRLSAERSARSHLLLOKQFYHSHRVQ
 PMALQVMSDEDEVDVADFEDROMLDFVDVNDKDEKQFMHLNMFYKORVIA
 DSHISWACAFSFEYKELHRYSLFMCWRFLIKLMNHGLVDSATINNCNTILENCR
 NSDDTTTNNNSVDRPSDSTNNNNIYDHPDINKNNVDKDNNSRDKVTK"
 853
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="a"
 1195
variation
 /gene="VRN2"
 /note="vrn2-1 mutation results in TGG Trp to TGA stop"
 /replace="a"
 1399
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="g"
 1487
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="a"
 1488
variation
 /gene="VRN2"
 /note="compared to the sequence in GenBank Accession
 Number AF180942"
 /replace="g"
 BASE COUNT 533 a 334 c 347 g 486 t
 ORIGIN
 Query Match 97.5%; Score 1678.6; DB 8; Length 1700;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1681; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

140 GACGTAGGGAATTCCTAATTTAGGGAGGCGCTCAGATGTTGCACCTACTTCAATTCGCG 199
 |||||||
 136 GACGTAGGGAATTCCTAATTTAGGGAGGCGCTCAGATGTTGCACCTACTTCAATTCGCG 195
 |||||||
 200 TCTTGACGTGTTGAGCTGTAATTTGAACAAGAAATGTGTAGGACAAATTTGGCGCGGAATC 259
 |||||||
 196 TCTTGACGTGTTGAGCTGTAATTTGAACAAGAAATGTGTAGGACAAATTTGGCGCGGAATC 255
 |||||||
 260 CTCACCGGAGGAAGATGTTCAACTGATGAAATCTCTTGATATATTTGAACCTGTGCG 319
 |||||||
 256 CTCACCGGAGGAAGATGTTCAACTGATGAAATCTCTTGATATATTTGAACCTGTGCG 315
 |||||||
 320 ACTATATTAACATCTTCACCTCGCTCTCTAGGCAACCCATGCTTCTTCCAAAGATGCTT 379
 |||||||
 316 ACTATATTAACATCTTCACCTCGCTCTCTAGGCAACCCATGCTTCTTCCAAAGATGCTT 375
 |||||||
 380 GAACCTCAAAAATTTGAGCAAAAGCGCAAAAAGAAAGTCAAGATCTACTGGAATGCTTTT 439
 |||||||
 376 GAACCTCAAAAATTTGAGCAAAAGCGCAAAAAGAAAGTCAAGATCTACTGGAATGCTTTT 435
 |||||||
 440 CAACCTTAAGGATGTAATTAACACATTAACAGAAACTGAAGTTAGGAGGAGATGTTCTTG 499
 |||||||
 436 CAACCTTAAGGATGTAATTAACACATTAACAGAAACTGAAGTTAGGAGGAGATGTTCTTG 495
 |||||||
 500 TCCATTTTCTCTATGCTATGCTGATGCTTCAAGGGGCTGCAATTTCAATTTGAATTCATC 559
 |||||||
 496 TCCATTTTCTCTATGCTATGCTGATGCTTCAAGGGGCTGCAATTTCAATTTGAATTCATC 555
 |||||||
 560 TCATGATTTATTTGAATTTGAGTTCAAGCTTTTGAAGAATACAGACAGTTAAATGTTTC 619
 |||||||
 556 TCATGATTTATTTGAATTTGAGTTCAAGCTTTTGAAGAATACAGACAGTTAAATGTTTC 615
 |||||||
 620 TGTAAACCTAATTCCTTCAATTTGAGGAAGAAGAAAGTGAATGATGATTAATTTGAGGC 679
 |||||||
 616 TGTAAACCTAATTCCTTCAATTTGAGGAAGAAGAAAGTGAATGATGATTAATTTGAGGC 675
 |||||||
 680 CTTCCTCTCTCTCGTGAACCTCGTAAAGCGGAGACAAAGAGTGGCAGAAATTAACACCG 739
 |||||||
 676 CTTCCTCTCTCTCGTGAACCTCGTAAAGCGGAGACAAAGAGTGGCAGAAATTAACACCG 735
 |||||||
 740 GAGACTTAAAGTATGCTTTTACCGTTGATTCACCCAGTTTAATTAATGCGACAGAGAA 799
 |||||||
 736 GAGACTTAAAGTATGCTTTTACCGTTGATTCACCCAGTTTAATTAATGCGACAGAGAA 795
 |||||||
 800 TGGAAATCACCTACTAATGATGGAAGACGCTGTTAGATATCCCGAGGCAACAGAGCT 859
 |||||||
 796 TGGAAATCACCTACTAATGATGGAAGACGCTGTTAGATATCCCGAGGCAACAGAGCT 855
 |||||||
 860 TCGTGACAAATTTGAGATGACGACGACAAATTCACCGACCATAGGCCACTTCTCTGGA 919
 |||||||
 856 TCGTGACAAATTTGAGATGACGACGACAAATTCACCGACCATAGGCCACTTCTCTGGA 915
 |||||||
 920 CGCTGCTGCTAAAGTTATTTGACAAAGCAAGCTGTGCTCCTCTACTTAAGCAAGAA 979
 |||||||
 916 CGCTGCTGCTAAAGTTATTTGACAAAGCAAGCTGTGCTCCTCTACTTAAGCAAGAA 975
 |||||||
 980 GTTATCTGTGAGCGCATGAGAGGCTAGAGCCACCTACTTCTTGAAGAACGCCAATTCCTA 1039
 |||||||
 976 GTTATCTGTGAGCGCATGAGAGGCTAGAGCCACCTACTTCTTGAAGAACGCCAATTCCTA 1035
 |||||||
 1040 TCATTTCTCAGACAGTCCACCAATGCGCTTGAAGCAATGATGTCGACCGGATAGCGA 1099
 |||||||
 1036 TCATTTCTCAGACAGTCCACCAATGCGCTTGAAGCAATGATGTCGACCGGATAGCGA 1095
 |||||||
 1100 GGATGAAGTCGATGACGATGTTGACGATTTTGAAGATGCCAGATGCTTGATGACTTTGT 1159
 |||||||
 1096 GGATGAAGTCGATGACGATGTTGACGATTTTGAAGATGCCAGATGCTTGATGACTTTGT 1155
 |||||||
 1160 GGATGAAGTCGATGACGATGTTGACGATTTTGAAGATGCCAGATGCTTGATGACTTTGT 1219
 |||||||
 1156 GGATGAAGTCGATGACGATGTTGACGATTTTGAAGATGCCAGATGCTTGATGACTTTGT 1215
 |||||||
 1220 AAGGCTTATAGCAGATGTCATATCTCTTGGGCAATGTGAAGCATTTTCAAGATTTAGCA 1279

DB	LOCUS	DEFINITION	ACCESSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	MODLINE	REFERENCE	TITLE	JOURNAL	FEATURES	source
Db	1216	AAAGGTTATAGCAGATGTGTCATTTTCGTTGGGCAATGTGAGCATTTTCAAGATTTTACGA	1275													
Qy	1280	GAAAGAGTTGCACCGTTACTCATCTCTTGTGGTGTGGAGATGTGTTTGTATTAACT	1339													
Db	1276	GAAAGAGTTGCACCGTTACTCATCTCTTGTGGTGTGGAGATGTGTTTGTATTAACT	1339													
Qy	1340	ATGGAACCTGAGCTGTTCGACTACGCCACCATCAACACTGCAATACCATCTCTGGAGA	1395													
Db	1336	ATGGAACCTGAGCTGTTCGACTACGCCACCATCAACACTGCAATACCATCTCTGGAGA	1395													
Qy	1400	TTGCGGTATAGCTTCAGACACACACACACCAACCAACCAACAGTGTGGATGTCCGAC	1459													
Db	1396	TTGCGGTATAGCTTCAGACACACACACACCAACCAACCAACAGTGTGGATGTCCGAC	1459													
Qy	1460	TCAGTCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1519													
Db	1456	TCAGTCAAAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1519													
Qy	1500	CAATGTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1515													
Db	1516	CAATGTGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC	1515													
Qy	1580	GCTTTTATGATCCGATTTATGCGATTTGATCTTCTCTTTTAAAAAATGTGTTA	1635													
Db	1576	GCTTTTATGATCCGATTTATGCGATTTGATCTTCTCTTTTAAAAAATGTGTTA	1635													
Qy	1640	GGAGCAACCAAAATTTTATGATGTTGATGTTCACTGATTAACATGTTTATGTTAAAAA	1695													
Db	1636	GGAGCAACCAAAATTTTATGATGTTGATGTTCACTGATTAACATGTTTATGTTAAAAA	1695													
Qy	1700	AAAAA 1704														
Db	1696	AAAAA 1700														
RESULT 5	AF284501	1722 bp mRNA	PLN 21-NOV-2001													
LOCUS	AF284501	Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete														
ACCESSION	AF284501															
KEYWORDS	AF284501.1 GI:16945789															
SOURCE																
ORGANISM		Arabidopsis thaliana														
		Arabidopsis thaliana														
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;														
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;														
		Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.														
		Gendall, A.R., Levy, Y., Wilson, A. and Dean, C.														
		The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of														
		Cell 107 (4), 525-535 (2001)														
		Verbalization in Arabidopsis														
		Cell 107 (4), 525-535 (2001)														
		2 (bases 1 to 1722)														

BASE COUNT	540 a	338 c	353 g	491 t
ORIGIN	<p> Query Match 95.6%; Score 1646.6; DB 8; Length 1722; Best Local Similarity 98.5%; Pred. No. 0; Matches 1681; Conservative 0; Mismatches 4; Indels 22; Gaps 1; </p>			
Db	<p> 20 TTGCTCTCTCTTACACAGCCATCGSTGTTTTCGAGATTTTCAGGCTTCATCCAAAGACA 79 16 TTGCTCTCTCTTACACAGCCATCGSTGTTTTCGAGATTTTCAGGCTTCATCCAAAGACA 75 80 TTCTATATTAAGCATTTGACAAAGAGGGGGTCTTAATTTGTCATTTAGTTTATCGCTAT 139 76 TTCTATATTAAGCATTTGACAAAGAGGGGGTCTTAATTTGTCATTTAGTTTATCGCTAT 135 140 GAGGTAGGAAATTTCTAATTTAGGGGAGGCTTCAGAGTTTGCACTAACTTCAATTCGGC 199 136 GAGGTAGGAAATTTCTAATTTAGGGGAGGCTTCAGAGTTTGCACTAACTTCAATTCGGC 195 200 TCTTGAGCTTGTAGGTATTTGAACAAGAAATGTTATGACGAATTTGCGCGGAAATC 259 196 TCTTGAGCTTGTAGGTATTTGAACAAGAAATGTTATGACGAATTTGCGCGGAAATC 255 260 CTCACGGAGAGATATTTCAACGTATGAGAAATCTTTGATATTTGTAACCTGTTGC 319 256 CTCACGGAGAGATATTTCAACGTATGAGAAATCTTTGATATTTGTAACCTGTTGC 315 320 ACTATATTAACATCTTTACCTCGCTCTCTTACAGCAACCATGTTTCTTCAAGATGCTT 379 316 ACTATATTAACATCTTTACCTCGCTCTCTTACAGCAACCATGTTTCTTCAAGATGCTT 375 380 GAACATCAAAATTTGAGCAAGAGCGGAAAGAAATCAAGATCTACTGGATGGATGTTT 439 376 GAACATCAAAATTTGAGCAAGAGCGGAAAGAAATCAAGATCTACTGGATGGATGTTT 435 440 CAACATTAAGATTTGATTAATACATTAACATTAACAGAAATCTGAAGTTAGGAGATTTGTTCTG 499 436 CAACATTAAGATTTGATTAATACATTAACATTAACAGAAATCTGAAGTTAGGAGATTTGTTCTG 495 500 TCATATTTTGCCTATGCTATGTGTGAGCTTCA-----AGGGC 537 496 TCATATTTTGCCTATGCTATGTGTGAGCTTCAAGGTGGGCAATTTCAACTGAGGGC 555 538 TGCATTTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTTCCAG 597 556 TGCATTTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTTCCAG 615 598 AATACCAAGACAGTTAATGTTTCTGTAAACAAATTTCTTCATATTTGAGAGAAGGAA 657 616 AATACCAAGACAGTTAATGTTTCTGTAAACAAATTTCTTCATATTTGAGAGAAGGAA 675 658 GTGATGACATTAATTTGAGGCTTCTCTCTCTGCTGAAACCTGTAAGCGAGACAA 717 676 GTGATGACATTAATTTGAGGCTTCTCTCTCTGCTGAAACCTGTAAGCGAGACAA 735 718 GAGGTGAGCAAAATTAACCCAGAGACTTAAAGTATGCTTTTACCGTGTGATTCACCA 777 736 GAGGTGAGCAAAATTAACCCAGAGACTTAAAGTATGCTTTTACCGTGTGATTCACCA 795 778 GTTAACTAATTTGAGCAAGAAATGATCACTTAAATATGAGAAACCGTGGTTAG 837 796 GTTAACTAATTTGAGCAAGAAATGATCACTTAAATATGAGAAACCGTGGTTAG 855 838 GATATCCCGAGCAAGAGCTTGTGAGCAATTTGAGATGACCGCAACTTCCACAG 897 856 GATATCCCGAGCAAGAGCTTGTGAGCAATTTGAGATGACCGCAACTTCCACAG 915 </p>			

QY	899	CATAGCCCACTCTCTCTGGACGCTGGGCTCAAAATTTATTTGCAACGCAAGCTGGGT	958
Db	901	CATAGCCCACTCTCTCTGGACGCTGGGCTCAAAATTTATTTGCAACGCAAGCTGTGT	960
QY	959	CCCTGCTACTAAGACAAGAAAGTTATCTGCTGACGANTCAGAGGCTAGAACCCACTACT	1018
Db	961	CCCTGCTACTAAGACAAGAAAGTTATCTGCTGACGANTCAGAGGCTAGAACCCACTACT	1020
QY	1019	TCCTTCAGAAAAGCCCAATTTCTATCTCTCACAGAGTCCAGCCAAATGGCGCTTGACCAAGT	1078
Db	1021	TCCTTCAGAAAAGCCCAATTTCTATCTCTCACAGAGTCCAGCCAAATGGCGCTTGACCAAGT	1080
QY	1079	AATGTCGACCGGGATTCGCGAGATGGAAGTCAGATGAGAGATGTCGAGATTTTGAAGATCG	1138
Db	1081	AATGTCGATTCGCGGATTCGCGAGATGGAAGTCAGATGAGAGATGTCGAGATTTTGAAGATCG	1140
QY	1139	CCAGATGCTTGATGACTTTTGATGATGTAATTAAGATGAAAAGCAATTCATGCATCTTGG	1198
Db	1141	CCAGATGCTTGATGACTTTTGATGATGTAATTAAGATGAAAAGCAATTCATGCATCTTGG	1200
QY	1199	GAACCTGCTTGTAAGAAAACAAAGGCTATATGAGATATGCTATCTTGGGATGTGA	1255
Db	1201	GAACCTGCTTGTAAGAAAACAAAGGCTATATGAGATATGCTATCTTGGGATGTGA	1260
QY	1259	AGCATTTTCAAGATTTTACGAGAAAGATGTCACCGTTACTATCACACTCTCTGGTGTG	1318
Db	1261	AGATTTTCAAGATTTTACGAGAAAGATGTCACACTGTTACTATCACACTCTCTGGTGTG	1320
QY	1319	GAGATTTGTTTTGATTTAACTATGGAACCATGACTGTGACTGACGTCAGCCACATCAACAA	1378
Db	1321	GAGATTTGTTTTGATTTAACTATGGAACCATGACTGTGACTGACGTCAGCCACATCAACAA	1380
QY	1379	CTGCAATACCATCTCGAGAAATGCGCTTAATAGTCTGACACCAACCAACCAACAA	1438
Db	1381	CTGCAATACCATCTCGAGAAATGCGCTTAATAGTCTGACACCAACCAACCAACAA	1440
QY	1439	CAACAGTGTGATTCGCTCCAGTGACTCAACACCAACCAATTAATTTGGATCATCC	1498
Db	1435	CAACAGTGTGATTCGCTCCAGTGACTCAACACCAACCAATTAATTTGGATCATCC	1499
QY	1499	CAATGACATTAACACAAGAACCAATGTTGACACACAGACATTAACAGCAGACAAAGT	1558
Db	1495	GAATGACATTAATTAACACAAGAACCAATGTTGACACACAGACATTAACAGCAGACAAAGT	1559
QY	1559	AATTAATTAAGAAAATCTCGGCTTTATGATACCGATTTTCGGATTTGTAACCTATCT	1618
Db	1554	AATTAATTAAGAAAATCTCGGCTTTATGATACCGATTTTCGGATTTGTAACCTATCT	1619
QY	1619	TCCTTTCTT-AAAAAATGTTTAGAGCAACCAATTTTATGTTAGTATTAATCACT	1677
Db	1614	TCCTTTCTTAAAAAATGTTTAGAGCAACCAAGATTTTATGTTAGTATTAATCACT	1678
QY	1678	GATTACATTTTGTAGTAAAAAAAAT 1705	
Db	1674	GATTACATTTTGTAGTAAAAAAAATGAT 1701	
RESULT 8			
AY034902			
LOCUS			
DEFINITION			
AY034902 1721 bp mRNA linear PLN 23-APR-2002			
ABRIDGED			
Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete cds.			
ACCESSION			
AY034902			
VERSION			
AY034902.1 GI:14334421			
SOURCE			
FLI CDNA			
ORGANISM			
Arabidopsis thaliana.			
Arabidopsis thaliana			
Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
Rosidae; eucosids II; Brassicales; Brassicaceae; Arabidopsis.			
1 (bases 1 to 1721)			
REFERENCE			
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,T., Chung,M.K.,			
AUTHORS			

TITLE	JOURNAL	REFERENCE
AUTHORS		
Goldsmith,A.D., Lee,J.M., Quach,H.L., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.	Arabidopsis full length cDNA clones	unpublished
2 (bases 1 to 1721)		
Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Narusaka,M., Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shinn,P., Southwick,A., Shinozaki,K., Davis,R.W., Ecker,J.R. and Theologis,A.	Direct Submission	
Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA		
RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-Length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.		
The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada,K., Liu,S.X., Sakano,H., Pham,P.K., Banh,J., Chung,M.K., Goldsmith,A.D., Lee,J.M., Quach,H.L., Tang,C.C., Toriumi,M., Yu,G., Bowser,L., Chen,H., Cheuk,R., Jones,T., Karlin-Neumann,G., Kim,C., Koeseema,E., Lam,B., Lin,J., Meyers,M.C., Miranda,M., Nguyen,M., Palm,C.J., Shinn,P., Southwick,A., Davis,R.W., Ecker,J.R. and Theologis,A.		
Yamada,K. (SSP/PGEC) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEC) contributed equally to this work as PIs.		
FEATURES		
source		
1. .1721		
/organism="Arabidopsis thaliana"		
/db_xref="taxon:3702"		
/chromosome="4"		
/clone="RAFL05-08-K19 (R09871)"		
/note="This clone is in a modified pluescript vector (Lambda ZAP) as a XhoI/SstI insert.		
ecotype: Columbia"		
1. .1721		
/gene="At4g16845"		
1. .231		
/gene="At4g16845"		
232. .1554		
/gene="At4g16845"		
/codon_start=1		
/evidence=experimental		
/product="unknown protein"		
/protein_id="AAK59409.1"		
/db_xref="GI:14334422"		
/translation="MCRONCRKSSPEEIVSTDENLILCYKPRVLYNFIHLRSGNPEFLEPLNFKIGAKRRKRSRSTGMYVFNKDNNTLQREVEDSCSPGSMCGSPFKRGKRGGRNTRLRKVCFLPLSPSLANGENCIALTLNNGENRLGIPETELTGQFEMTSNIPALNHSLSIDAGAKVILTTEAVVAPATKTRKLSAENSEASHLLQRPQYHSRVAQPMALQYVMSDRSDSEVDVADFEQMDLDDFYDVNDEKQFHLNWSFYRKQRYVADGHSVMACEVFSRYEKEILHCYSLSLFCWRLFLIKLNHGLVDSATINCMTILENCFNSTGNTNNNSVHDPSDSNTNNNNIVDHPIDIKKNVNDKNNDSRDK"		
1555. .1721		
/gene="At4g16845"		
BASE COUNT	548 a	324 c 353 g 496 t
ORIGIN		
Query Match	91.1%	Score 1568.6; DB 8; Length 1721;
Best Local Similarity	96.5%	Pred. No. 0;
Matches 1648; Conservative	0;	Mismatches 49; Indels 10; Gaps 4;

QY 2 AACCTTCTCAATTTTGGTTG--CTCTCTCTTACACAGCAATCGGTGTTTCCAGCTT 59
 Db 1 AAGCTTCTCAATTTTGGTTGCTCTCTCTTACACAGCAATCGGTGTTTCCAGCTT 60
 QY 60 TCAGGCTCAATTCAGACATCTTATATAGCATTTTCAGAGGCGGTTCTAATGT 119
 Db 61 TCAGGCTCAATTCAGACATCTTATATAGCATTTTCAGAGGCGGTTCTAATGT 120
 QY 120 TGCATGAGTTTATGCTTATGAGTGAAGAAATTTATATAGGAGGCTTCAGATT 179
 Db 121 TGCATGAGTTTATGAGTGAAGTGAAGAAATTTATATAGGAGGCTTCAGATT 180
 QY 180 GCCTTACCTTCAATATGAGTGAAGTGAAGAAATTTATATAGGAGGCTTCAGATT 239
 Db 181 GCCTTACCTTCAATATGAGTGAAGTGAAGAAATTTATATAGGAGGCTTCAGATT 240
 QY 240 CAGAAATTTGCGGCAATTCCTACCGGAGAGTGAATTCACATGATGAGAAATCTTTG 299
 Db 241 CAGAAATTTGCGGCAATTCCTACCGGAGAGTGAATTCACATGATGAGAAATCTTTG 300
 QY 300 ATATATTTGTAACCTGTTGACTATATATACATCTTTCACCTGCTCTGAGCAACCA 359
 Db 301 ATATATTTGTAACCTGTTGACTATATATACATCTTTCACCTGCTCTGAGCAACCA 360
 QY 360 TCGTTTCTTCAAGATGCTTGAATCAACAAATTTGAGCAAGCGCAAAAGAAATTCAGA 419
 Db 361 TCGTTTCTTCAAGATGCTTGAATCAACAAATTTGAGCAAGCGCAAAAGAAATTCAGA 420
 QY 420 TCTACTGGATGGTGTGTTTCAACTATATAGATTTATATACATTCAGAAATCTGAA 479
 Db 421 TCTACTGGATGGTGTGTTTCAACTATATAGATTTATATACATTCAGAAATCTGAA 480
 QY 480 GTTAGGAGATTTGTTCTGCTTATTTGCTTATGCTATGCTATGCTTCAAGGCTG 539
 Db 481 GTTAGGAGATTTGTTCTGCTTATTTGCTTATGCTATGCTATGCTTCAAGGCTG 540
 QY 540 CAATTCATTTGATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTGATTTG 599
 Db 541 CAATTCATTTGATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTGATTTG 600
 QY 600 TACCAAGAGTTAATGTTTCTGTAAACTTAATTCCTTCAATTTGAGAGAAAGAGT 659
 Db 601 TACCAAGAGTTAATGTTTCTGTAAACTTAATTCCTTCAATTTGAGAGAAAGAGT 660
 QY 660 GATGAGATTAATTTGAGGCTTCTCTCTCTGCTGGAACCTCGTAAGCGGAGCAAGA 719
 Db 661 GATGAGATTAATTTGAGGCTTCTCTCTCTGCTGGAACCTCGTAAGCGGAGCAAGA 720
 QY 720 GGTGGAGAAATTAACCCGAGGACTTAAGATGCTTTTACGTTGGATTCACCGAGT 779
 Db 721 GGTGGAGAAATTAACCCGAGGACTTAAGATGCTTTTACGTTGGATTCACCGAGT 780
 QY 780 TTACTTAATGCGACAGAAATGGAATCAACCTTAAATGATGGAACCGTGTGTAAGA 839
 Db 781 TTACTTAATGCGACAGAAATGGAATGCAACCTTAAATGATGGAACCGTGTGTAAGA 840
 QY 840 TATCCGAGGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTCACCAAGCC 899
 Db 841 TATCCGAGGCAACAGAGCTTCTGAGCAATTTGAGATGACCAACATTCACCAAGCC 900
 QY 900 ATAGCCCACTTCTCTGAGGCTGCTGCTAAGTTATGATGAGCAAGGAAAGCTGTGTC 959
 Db 901 ATAGCCCACTTCTCTGAGGCTGCTGCTAAGTTATGATGAGCAAGGAAAGCTGTGTC 960
 QY 960 CCGCTACTTAAGCAAGAAATGATCTGCTGAGGATTCAGAGGCTTAAGAGGCACTTACT 1019
 Db 961 CCGCTACTTAAGCAAGAAATGATCTGCTGAGGATTCAGAGGCTTAAGAGGCACTTACT 1020
 QY 1020 CTTGAGAAAGCCAAATTTATCATTTCTCAGAGTTCAGCCAAATGGCGCTTGAGCAAGTA 1079
 Db 1021 CTTGAGAAAGCCAAATTTATCATTTCTCAGAGTTCAGCCAAATGGCGCTTGAGCAAGTA 1080

QY 1080 ATGCTGACCGGATAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
 Db 1081 ATGCTGATCGGATAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1140 CAGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1199
 Db 1141 CAGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 QY 1200 AACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
 Db 1201 AACTGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1260
 QY 1260 GCATTTTCAAGATTTTACAGAAAGTTCACCGCTTACTCATCTCTTCTGCTGTG 1319
 Db 1261 GCATTTTCAAGATTTTACAGAAAGTTCACCGCTTACTCATCTCTTCTGCTGTG 1320
 QY 1320 AGATTTGTTTGTATTAACATTTGAGAACATGAGCTTGTGACTCAGCCACATCAACAC 1379
 Db 1321 AGATTTGTTTGTATTAACATTTGAGAACATGAGCTTGTGACTCAGCCACATCAACAC 1380
 QY 1380 TGCATTAACATCTCGAGAAATTTGCCGTAATAGCTCAGACACACACCAACCAACAC 1439
 Db 1381 TGCATTAACATCTCGAGAAATTTGCCGTAATAGCTCAGACACACACCAACCAACAC 1434
 QY 1440 AACAGTGTGATGCTGCTGAGTCAACAGCAACCAACCAACCAACCAACCAACCAACCAAC 1499
 Db 1435 AACAGTGTGATGCTGCTGAGTCAACAGCAACCAACCAACCAACCAACCAACCAACCAAC 1494
 QY 1500 AATGCAATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1559
 Db 1495 AATGCAATTAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1553
 QY 1560 AATTAATGAGAAATCTCGGCTTTATGATTAACCAATTTATGATGATGATGATGATG 1619
 Db 1554 AATTAATGAGAAATCTCGGCTTTATGATTAACCAATTTATGATGATGATGATGATG 1613
 QY 1620 CTTTCTT-AAAGAAATGTTAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1678
 Db 1614 CTTTCTTAAAAAAAAATGTTAGAGCAACCAACCAACCAACCAACCAACCAACCAACCAAC 1673
 QY 1679 ATTACATTTTATGTTAAAAAATAAT 1705
 Db 1674 ATTACATTTTATGTTAAAAAATAAT 1700

RESULT 9
 AX032894 1737 bp DNA linear PAT 21-SEP-2000
 LOCUS Sequence 7 from Patent WO0044918.
 DEFINITION AX032894
 ACCESSION AX032894.1 GI:10279829
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 1 (bases 1 to 1737)
 AUTHORS Dean, C. and Gendall, A.
 TITLE Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 JOURNAL Patent: WO 0044918-A 7 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)

FEATURES
 source location/Qualifiers
 1..1737
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 543 a 330 c 359 g 505 t
 ORIGIN
 Query Match 89.3%; Score 1537.6; DB 6; Length 1737;
 Best Local Similarity 95.3%; Pred. No. 0;

Matches 1649; Conservative 0; Mismatches 49; Indels 32; Gaps 5;

```

OY 1 CAAGCTTCTTCATTTTGGCTG--CTCTCTTACACGCAATGSGTGTTCGAGCT 58
Db 1 CAAGCTTCTTCATTTTGGCTGCTCTCTCTTACACGCGCAATCGTGTTCGAGCT 60
OY 59 TTCAGGCTTCATTCACAGATTTATATATAGCATATTCAGAGAGGCGGTTCATTTG 118
Db 61 TTCAGGCTTCATTCACAGATTTATATATAGCATATTCAGAGAGGCGGTTCATTTG 120
OY 119 TTGCATGTAGTTTATGCTATGACGTAGGAAAATTCATAATTAAGGGAGCGCTCAGATT 178
Db 121 TTGCATGTAGTTTATGCAATATGACGTAGGAAAATTCATAATTAAGGGAGCGCTCAGAGTT 180
OY 179 TGCATTAATCTATATGCGCTTTCAGCTGTGTGCTATTAATGAAACAGAAATGCTAG 238
Db 181 TGCATTAATCTATATGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 240
OY 239 GCAGAAATGTCGCGGAAATCCTCAACGAGAGAGATTTTCAGCTGATGAGATCTCTT 298
Db 241 GCAGAAATGTCGCGGAAATCCTCAACGAGAGAGATTTTCAGCTGATGAGATCTCTT 300
OY 299 GATATATTTGTAACCTGTTGCACTATATACATCTTTGCTGCTCTGAGCAACCC 358
Db 301 GATATATTTGTAACCTGTTGCACTATATACATCTTTGCTGCTCTGAGCAACCC 360
OY 359 ATCGTTTCTTCCAAAGATGCTTACATACAAATTTGGAGCAAGCGCAAGAAAGATCAAG 418
Db 361 ATCGTTTCTTCCAAAGATGCTTACATACAAATTTGGAGCAAGCGCAAGAAAGATCAAG 420
OY 419 ATCTACTGGGATGTAGTTTTCAGTATATAGATTTGTAATTAACATTTACAGAAACCTGA 478
Db 421 ATCTACTGGGATGTAGTTTTCAGTATATAGATTTGTAATTAACATTTACAGAAACCTGA 480
OY 479 ACTTGGAGAGATTTGTTCTTCTGCAATTTGCTCTATGCTATGCTGAGCTTCA----- 531
Db 481 ACTTGGAGAGATTTGTTCTTCTGCAATTTGCTCTATGCTATGCTGAGCTTCAAGGTGG 540
OY 532 -----AGGGGCTGCAATTTGTAATTCATCTCATCTCATGTTATTTGAT 576
Db 541 CAACATTTACAACTGAGGGCTGCAATTTGTAATTCATCTCATCTCATGTTATTTGAT 600
OY 577 TTGAGTTCAAGCTTTTCGAGAAATACACAGAGATTAATGTTTCTGTAACCTTAATTCCT 636
Db 601 TTGAGTTCAAGCTTTTCGAGAAATACACAGAGATTAATGTTTCTGTAACCTTAATTCCT 660
OY 637 TCATATTTGAGAAAGAGATGATGACGATTAATTTGAGCCCTCTCTCTGCTGCA 696
Db 661 TCATATTTGAGAAAGAGATGATGATGATTAATTTGAGCCCTCTCTCTGCTGCA 720
OY 697 AACCTGTAAGCGAGACAAAGAGGAGCAAAATTAACAGAGAGACTTAAGTATGCT 756
Db 721 AACCTGTAAGCGTAAGCAAAAGAGGAGCAAAATTAACAGAGAGACTTAAGTATGCT 780
OY 757 TTTTACCGTTGATTCACCCAGTTTAATGAGCAGAGAAATGGAATCACCTACTTA 816
Db 781 TTTTACCGTTGATTCACCCAGTTTAATGAGCAGAGAAATGGAATGCTCCTGCTGA 840
OY 817 ANGATGGAACCGGTGTTAGATATCCGAGCAACAGAGCTGTGTCAGCAATTTGACA 876
Db 841 ANGATGGAACCGGTGTTAGATATCCGAGCAACAGAGCTGTGTCAGCAATTTGACA 900
OY 877 TGACGAGCAACATTCACACAGCCATAGCCCACTTCTCTGAGCGGTGCTTAAGTTA 936
Db 901 TGACGAGCAACATTCACACAGCCATAGCCCACTTCTCTGAGCGGTGCTTAAGTTA 960
OY 937 TATTAGCAAGCAAGCTGTGCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGAT 996
Db 961 TATTAGCAAGCAAGCTGTGCTGCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGAT 1020
OY 997 CAGAGGCTTAGAGCCACACTACTTCTTCAGAAACGCCAATTCATATTCATTCACAGAGTCC 1056
Db 1021 CAGAGGCTTAGAGCCACACTACTTCTTCAGAAACGCCAATTCATATTCATTCACAGAGTCC 1080

```

```

OY 1057 AGCCAAATGGCGCTTGAGCAAGTATGCTGACCGGATAGCCAGATGAAGTCGATGAGC 1116
Db 1081 AGCCAAATGGCGCTTGAGCAAGTATGCTGATCGGATAGCCAGATGAAGTCGATGAGC 1140
OY 1117 ANGTTGCAATTTTGAAGATTCGCCAGATGCTGATGACTTGTGATGTAATTAAGATG 1176
Db 1141 ANGTTGCAATTTTGAAGATTCGCCAGATGCTGATGACTTGTGATGTAATTAAGATG 1200
OY 1177 AAAGCAATTCATGCAATCTTTGGAAGTCTGTTGTAAGAAACAAAGGTTATAGCAGATG 1236
Db 1201 AAAGCAATTCATGCAATCTTTGGAAGTCTGTTGTAAGAAACAAAGGTTATAGCAGATG 1260
OY 1237 GTCATATCTCTTGGGATGATGAAGATTTTCAAGATTTTACAGGAAGAGTTGACCGCTT 1296
Db 1261 GTCATATCTCTTGGGATGATGAAGATTTTCAAGATTTTACAGGAAGAGTTGACCGCTT 1320
OY 1297 ACTCATCACTCTTCTGCTGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1356
Db 1321 ACTCATCACTCTTCTGCTGAGATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1380
OY 1357 TCGACTCAGCCACATCAACACCTGCAATACATCTCTGAGATTTGCCGTTATAGCTCAG 1416
Db 1381 TCGACTCAGCCACATCAACACCTGCAATACATCTCTGAGATTTGCCGTTATAGCTCAG 1440
OY 1417 ACACGACACACACACACACACACACACACACACACACACACACACACACACACACAC 1476
Db 1441 -----TCACTAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 1494
OY 1477 ACAATTAACATTTGTGATCATCCCAATGACATTAACAACAACAACAACAACAACAACA 1536
Db 1495 ACAATTAACATTTGTGATCATCCCAATGACATTAACAACAACAACAACAACAACAACA 1554
OY 1537 ACAATTAACGACAGACAGCAAGATTAATTAAGAAATATCCGCTTTATGATACCGAT 1596
Db 1555 ACAATTAACGACAGAGAC- AAGTAAATTAATAGAAACACCTCGGTTAATATATACGAT 1613
OY 1597 TTATCGATTTGATCACTATCTCTCTCTCTT-AAAAATGTTAGAGCAACAATTTT 1655
Db 1614 CTATCGATTTGATCACTATCTCTCTCTCTTCTTTAAAAAATGTTAGAGCAACAAGATTT 1673
OY 1656 TTATATGTTAGTATGATCACTGATTAATTTTATGTTAAAAAATTTT 1705
Db 1674 TTATTTGTTAGTATGATCACTGATTAATTTTATGTTAAAAAATTTTATGAT 1723

RESULT 10
AY063047 1354 bp mRNA linear PLN 18-APR-2002
LOCUS Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete
DEFINITION
cds.
ACCESSION AY063047
VERSION AY063047.1 GI:17104664
KEYWORDS FLI_CDNA.
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
REFERENCE 1 (bases 1 to 1354)
Yamada, K., Liu, S. X., Sakano, H., Pham, P. K., Banh, J., Etgu, P.,
Lee, J. M., Toriumi, M., Yu, G., Brooks, S., Chao, Q., Chen, H.,
Karlin-Neumann, G., Kim, C., Lam, B., Miranda, M., Nguyen, M.,
Palm, C. J., Shinn, P., Southwick, A., Davis, R. W., Ecker, J. R. and
Theologis, A.
Arabidopsis Open Reading Frame (ORF) Clones
Unpublished
2 (bases 1 to 1354)
Yamada, K., Banh, J., Banno, F., Chang, E., Dale, J. M., Goldsmith, A. D.,
Lee, J. M., Onodera, C. S., Quach, H. L., Tang, C. C., Toriumi, M., Wu, H. C.,
Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H.,
Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamuya, A.,
Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Liu, J.,

```


TITLE
JOURNAL

COMMENT

Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A., Direct Submission Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street Albany, CA 94710, USA

The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : RIKEN Arabidopsis Full-length cDNA) : Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamuya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the PUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Baah, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shim, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES

source

Location/Qualifiers

1..1354

/organism="Arabidopsis thaliana"

/db_xref="taxon:3702"

/chromosome="4"

/clone="U09871"

/note="this clone is in PUNI 51."

ecotype: Columbia"

1..1354

/gene="At4g16845"

1..1323

/gene="At4g16845"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AA134221.1"

/db_xref="GI:17104655"

/translation="MCRONCRRAKSSPEVYSTDENLLIYCKFVRILNIFHLRLNPS

FLPRLNLTGAKRRKRSSTGIVNFKCNNTLQRTVREDSCPSCLSGSFRK

LQFLNNSHDLFEFEKLEEDYGVNFKCNNTLQRTVREDSCPSCLSGSFRK

RQGRNTRRLKVCFLPDLSPISLANGTEGIALINDGRGLGPEALRLEAGQEMTS

NIPIPAHSSLDAGAKVILITTEAVVPATKRLSARSSEARSHLLOKROFESHRYO

PMALQVNSDRSDSEVDVADFEEDROMLDFVKNDEKOFMLHNSFVKORFVIA

DGHISWACEVRSFEKELHCYSLEFCWMLFLIKLNHGLVDSATINNCTILENCR

NTSYNNNNNSVDHPSDSNTNNNNIYDHPIDIKNNNNVDKNNNSRDK"

1324..1354

/gene="At4g16845"

BASE COUNT 431 a 264 c 287 g 372 t

ORIGIN

Query Match 74.2% Score 1277.4; DB 8; Length 1354;
Best Local Similarity 97.2%; Pred No. 2.4e-272;
Matches 1333; Conservative 0; Mismatches 31; Indels 7; Gaps 2;

QY 231 ATGTGTAGGACGATTTGCGCGCAAAATCTCACCGAGAGAGTTCACACTGATGAG 290
DB 1 ATGTGTAGGACGAAATGTGCGCGCAAAATCTCACCGAGAGAGTTCACACTGATGAG 60
QY 291 AATCTCTGATATATATGTAACCTGTTCGACTATATTAACATCTTTCACCTGCTCTTA 350
DB 61 AATCTCTGATATATATGTAACCTGTTCGACTATATTAACATCTTTCACCTGCTCTTA 120
QY 351 GGCAACCATGTTCTTCCAAAGATGCTTGAACATAAATTTGAGCAAGGCAAAAGA 410
DB 121 GGCAACCATGTTCTTCCAAAGATGCTTGAACATAAATTTGAGGCAAGGCAAAAGA 180
QY 411 AAGTCAAGTCTACGGGATGAGTTCATCACTATTAAGATGATGATTAACATTAAG 470
DB 181 AAGTCAAGTCTACGGGATGAGTTCATCACTATTAAGATGATGATTAACATTAAG 240

QY 471 AAACCTGAAGTGAAGGAGAGATTTGTTCTTGCATATTTGCTATGCTATGCTGACCTTC 530
DB 241 AGACCTGAAGTGAAGGAGAGATTTGTTCTTGCATATTTGCTATGCTATGCTGACCTTC 300
QY 531 AAGGGCTGCAATTTCAATTTGAATTCATTCATGATTAATTTGAATTTGAATTTGAATTT 590
DB 301 AAGGGCTGCAATTTCAATTTGAATTCATTCATGATTAATTTGAATTTGAATTTGAATTT 360
QY 591 TTGGAAGATTAACAGACATTAATGTTCTGTAAACCTTAATTCCTTCAATTTGAGAA 650
DB 361 TTGGAAGATTAACAGACATTAATGTTCTGTAAACCTTAATTCCTTCAATTTGAGAA 420
QY 651 GAAGGAAGTGAAGGAGATTAATTTGAGGCTTCTCTCTGCTGAAACCTGTAAGCG 710
DB 421 GAAGGAAGTGAAGGAGATTAATTTGAGGCTTCTCTCTGCTGAAACCTGTAAGCGT 480
QY 711 AGCAAAAGAGTGAAGGAGATTAACAGAGACTTAAGTATGCTTTTACCGTGGAT 770
DB 481 AGCAAAAGAGTGAAGGAGATTAACAGAGACTTAAGTATGCTTTTACCGTGGAT 540
QY 771 TCACCCAGTTTAATTAATGCAAGAAATGGAATCACTACTTAATGATGAACCGT 830
DB 541 TCACCCAGTTTAATTAATGCAAGAAATGGAATCACTACTTAATGATGAACCGT 600
QY 831 GGTTAAGATATCCCGAGGCAACAGAGCTTCTGACAAATTTGATGACCGCAACAT 890
DB 601 GGTTAAGATATCCCGAGGCAACAGAGCTTCTGACAAATTTGATGACCGCAACAT 660
QY 891 CCACGAGCATGAGCCACTCTTCTCTGAGCGCTGCTGCTAAATTTGATGACCGCAAC 950
DB 661 CCACGAGCATGAGCCACTCTTCTCTGAGCGCTGCTGCTAAATTTGATGACCGCAAC 720
QY 951 GCGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1010
DB 721 GCGTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
QY 1011 CACCTACTTCTTCAAAAGCCCAATTCATCTCTGACAGATCCAGCCATGCGCTT 1070
DB 781 CACCTACTTCTTCAAAAGCCCAATTCATCTCTGACAGATCCAGCCATGCGCTT 840
QY 1071 GAGCAAGTATGCTGACCGGAGATGAGAGATGAATGATGATGATGATGATGATGATG 1130
DB 841 GAGCAAGTATGCTGACCGGAGATGAGAGATGAATGATGATGATGATGATGATGATG 900
QY 1131 GAAGATCGCAGATGCTGATGACTTTGTGATGTGAATTAAGTGAAGAAAGCAATTCATG 1190
DB 901 GAAGATCGCAGATGCTGATGACTTTGTGATGTGAATTAAGTGAAGAAAGCAATTCATG 960
QY 1191 CATCTTTGGAATGCTTTGTAAGAAACAAAGGTTTATGACATGCTATGCTTGG 1250
DB 961 CATCTTTGGAATGCTTTGTAAGAAACAAAGGTTTATGACATGCTATGCTTGG 1020
QY 1251 GCATGTGAAGCTTTTCAAGATTTTGAAGAAAGTGTGACCGTATGACATGCTTTC 1310
DB 1021 GCATGTGAAGCTTTTCAAGATTTTGAAGAAAGTGTGACCGTATGACATGCTTTC 1080
QY 1311 TGGTGTGAGATTTTGTGATTAATTAATGAAGCAACCATGACTTGTGACTGACGAC 1370
DB 1081 TGGTGTGAGATTTTGTGATTAATTAATGAAGCAACCATGACTTGTGACTGACGAC 1140
QY 1371 ATCAACAACTGCAATACCATCTCGAAGAAATGCGGTAAATGCTGACACACCAAC 1430
DB 1141 ATCAACAACTGCAATACCATCTCGAAGAAATGCGGTAAATGCTGACACACCAAC 1194
QY 1431 AACCAACAAACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1490
DB 1195 AACCAACAAACAGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1254
QY 1491 GATCATCCCAATGATTAATTAATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1550
DB 1255 GATCATCCCAATGATTAATTAATGAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 1314

[illegible]

RESULT 13	AX032893	5895 bp	DNA	Linear	PAT 21-SEP-2000
LOCUS	AX032893				
DEFINITION	Sequence 6 from Patent WO0044918.				
ACCESSION	AX032893				
VERSION	AX032893.1				
KEYWORDS					
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 5895)				
AUTHORS	Dean, C. and Gendall, A.				
TITLE	Methods and means for modification of plant characteristics using the vernalization gene vrn2				
JOURNAL	Patent: WO 0044918-A 6 03-AUG-2000;				
	DEAN CAROLINE (SB) ; GENDALL ANTHONY (SB) ; PLANT BIOSCIENCE LTD (GB)				
FEATURES					
SOURCE	Location/Qualifiers				
	1..5895				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
BASE COUNT	1742 a 1027 c 1070 g 2056 t				
ORIGIN					
Query Match	18.5%; Score 319.4; DB 6; Length 5895;				
Best Local Similarity	92.7%; Pred. No. 3.1e-60;				
Matches 370; Conservative	0; Mismatches 21; Indels 8; Gaps 3;				
QY	1308 TTCTGGGTGGAGATGTTTTCATTAATTAAGAAACCATGGACCTTGTGCATCAGCC	1367			
DB	4677 TCCAGGTGTGGAGATTGTTTTCATTAATTAAGAAACCATGGACCTTGTGCATCAGCC	4736			
QY	1368 ACCATCAACAACCTGCATACCATCTCGAGAAATGCCGTATAGCTCAGACACACACC	1427			
DB	4737 ACCATCAACAACCTGCATACCATCTCGAGAAATGCCGTATAGCTCAG-----TCACT	4750			
QY	1428 ACCAACAACAACAACAGTGTGATGTCGCCAGTGTCAACACCAACAACATACATT	1487			
DB	4791 AACCAACAACAACAACAGTGTGATGTCGCCAGTGTCAACACCAACAACATACATT	4850			
QY	1488 GTGATATCCCATGATGATTAACAACAAGAAATGTGCACACAGAAATTAACAGC	1547			
DB	4851 GTGATATCCCATGATGATTAACAACAAGAAATGTGCACACAGAAATTAACAGC	4910			
QY	1548 AGAACAACAAGTAATTAATAGAAATCTCCGCTTTATGATACCGATTATCGGATTG	1607			
DB	4911 AGAACAACAAGTAATTAATAGAAATCTCCGCTTTATGATACCGATTATCGGATTG	4969			
QY	1608 TAACTATATCTCTCTCTT-AAAAAATGTTTAGAGCAAAATTTTATATGTTAG	1666			
DB	4970 TAACTATATCTCTCTCTT-AAAAAATGTTTAGAGCAAAATTTTATATGTTAG	5029			
QY	1667 TGTATTCACCTGATTACATTTTATAGTTAAAAAAAAT 1705				
DB	5030 TGTATTCACCTGATTACATTTTATAGTTAAAAAAAATGAT 5068				
RESULT 14					
LOCUS	ATCHRIV45				
DEFINITION	Arabidopsis thaliana DNA chromosome 4, contig fragment No. 45.				
ACCESSION	AL161545				
VERSION	AL161545.2				
KEYWORDS	GI:7268431				
SOURCE	Arabidopsis thaliana.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eucotsids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 199548)				
AUTHORS	EU Arabidopsis sequencing project.				

[illegible]

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 17:43:25 ; Search time 422 Seconds

(without alignments)
9189.434 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722
Sequence: 1 caagcttctcaatttctc.....aatgattctcttact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq.101002.*

1:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24:	/SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1722	100.0	1722	21	AAA63670
2	1722	100.0	1722	21	AAA47751
3	1569.6	91.1	1715	21	AAA63671
4	1569.6	91.1	1715	21	AAA47753
5	1537.6	89.3	1737	21	AAA47755
6	979.8	56.9	1497	21	AAAC51581
7	412.2	23.9	6338	21	AAA47752
8	319.4	18.5	5895	21	AAA47754
9	314.2	18.2	319	21	AAAC51582

10	306.4	17.8	640	21	AAAC44147	Arabidopsis thaliana
11	291.4	16.9	2280	21	AAAC36947	CDNA encoding a MP
12	252.4	14.7	2248	21	AAAC36953	CDNA encoding an O
13	230.8	13.4	558	21	AAAC47758	At Hyp 2245035 (AT
14	170.8	9.9	525	21	AAAC54919	Arabidopsis thaliana
15	163.8	9.5	325	21	AAAC47757	At163743 Expressed
16	126.6	7.4	463	21	AAAC47756	Rice expressed seq
17	71	4.1	8045	21	AAAC5171	Neurospora crassa
18	69.6	4.0	567	21	AAAC29550	HIV codon altered
19	69	4.0	198	17	AAAT17204	DNA-spanner oligon
20	69	4.0	198	17	AAAT17205	DNA-spanner oligon
21	65.8	3.8	198	17	AAAT17202	DNA-spanner oligon
22	65.8	3.8	198	17	AAAT17203	DNA-spanner oligon
23	63.2	3.7	1216	21	AAAC36962	DNA fragment encod
24	62.2	3.6	5580	21	AAAC36948	CDNA encoding a MPC
25	62	3.6	5511	21	AAAC61847	Cryptosporidium pa
26	62	3.6	5511	24	ABT04776	C parvum GP900 gen
27	62	3.6	7334	21	AAAC61846	Cryptosporidium pa
28	62	3.6	7334	24	ABT04775	C parvum GP900 gen
29	60.4	3.5	5163	19	AAAC20700	Cryptosporidium pa
30	60.4	3.5	5163	21	AAAC61849	ORF encoding a por
31	60.4	3.5	5163	24	ABT04778	C parvum GP900 gen
32	60.4	3.5	5318	19	AAAC20701	Cryptosporidium pa
33	60.4	3.5	5318	21	AAAC61848	DNA encoding a por
34	60.4	3.5	5318	24	ABT04777	C parvum GP900 gen
35	59	3.4	61020	22	AAAC61788	Tumour suppressor
36	56.8	3.3	2336	23	ABT25662	Drosophila melanog
37	55.8	3.2	3300	21	AAAC55699	DNA encoding yeast
38	54.6	3.2	4829	23	ABT13399	Drosophila melanog
39	54.6	3.2	39746	23	ABT13398	Drosophila melanog
40	54.2	3.1	8524	22	AAAC6588	Tumour suppressor
41	54.2	3.1	8524	24	ABT33841	Human immune syste
42	54	3.1	144	12	AAQ12515	CSP-2 peptide from
43	53.4	3.1	1690	16	AAQ92524	Nicotiana glauca se
44	53.2	3.1	2892	24	ABT79885	Fungal ZBC gene se
45	52.2	3.0	8662	24	ABT34637	Human metastasis a

ALIGNMENTS

RESULT 1

ID	AAA63670	standard; cDNA; 1722 BP.
AC	AAA63670;	
XX		
DT	04-DEC-2000 (first entry)	
XX		
DE	Nucleotide sequence of VRN2 sequence of Arabidopsis sp.	
XX		
KW	H51; one locus-FRIGIDA; FRI gene; flowering time; blotting;	
KW	flower initiation; stem elongation; flower production; VRN2; ss.	
XX		
OS	Arabidopsis sp.	
XX		
PN	WO200046358-A2.	
XX		
PD	10-AUG-2000.	
XX		
PF	25-JAN-2000; 2000WO-GB00197.	
XX		
PR	05-FEB-1999; 99GB-0002660.	
XX		
XX	(PLAN-) PLANT BIOSCIENCE LTD.	
XX	Johanson U, West J, Dean C;	
XX	WPI; 2000-532899/48.	
XX	New nucleic acid derived from the FRI locus of a plant, e.g.	
PT	Arabidopsis, encoding a polypeptide capable of specifically altering	
PT	the flowering time of a plant	

DE VRN2 nucleic acid.

XX Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ss.

XX Arabidopsis thaliana var 'Landsberg erecta'.

OS

XX Key location/qualifiers
 FT 231..1568
 FT CDS /tag- a
 FT /product- "VRN2 polypeptide"
 FT replace(119,8)
 FT mutation /tag- b
 FT /note- "found in mutant vrn2-1; alters a Trp codon
 FT to a stop codon leading to a truncation of
 FT the protein"

XX WO200044918-A1.

XX PD 03-AUG-2000.

XX 28-JAN-2000; 2000WO-GB00248.

XX PR 28-JAN-1999; 99GB-0001927.

XX PA (PLAN-) PLANT BIOSCIENCE LTD.

XX Dean C. Gendall A;

XX PFI: 2000-499333/44.

XX DR P-PSDB; AAB00060.

XX Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success

XX Claim 4: Fig 6; 105pp; English.

XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.

XX Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

SO

Query Match 100.0%; Score 1722; DB 21; Length 1722;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAAGCTTCTTCGAATTTGCTGCTCTCTTACACAGCCAAATCGGTGTTTGCAGCTTT 60
 DB 1 CAAGCTTCTTCGAATTTGCTGCTCTCTTACACAGCCAAATCGGTGTTTGCAGCTTT 60
 OY 61 CAGGCTCAATCCAGACATCTTATATAGCATATGAGAAAGGCGGCTTCTAATTTT 120
 DB 61 CAGGCTCAATCCAGACATCTTATATAGCATATGAGAAAGGCGGCTTCTAATTTT 120
 OY 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180
 DB 121 GCATTGAGTTATGCTATGACGTAGGGAATTTCAATTTAGGGAGGCTCAGAGTTTG 180
 OY 181 CACTTAATCTCAATTCGGCTCTTGACGTTGTTGAGTGAATGAACAAGAAATGTGTAGGC 240
 DB 181 CACTTAATCTCAATTCGGCTCTTGACGTTGTTGAGTGAATGAACAAGAAATGTGTAGGC 240
 OY 241 AGAATGTGCGGCAAAATCTCACCAGGAGAAAGTGAATTTCAACTGATGAGATCTCTTGA 300
 DB 241 AGAATGTGCGGCAAAATCTCACCAGGAGAAAGTGAATTTCAACTGATGAGATCTCTTGA 300

OY 301 TATATTGTAACCGTTGACATATATAACATCTTACACTTGGCTCTGAGCAACCCAT 360
 DB 301 TATATTGTAACCGTTGACATATATAACATCTTACACTTGGCTCTGAGCAACCCAT 360
 OY 361 CGTTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGGCAAAAGAAAGTCAGAT 420
 DB 361 CGTTTCTTCCAAAGATGCTTGAACCTACAAATTTGAGCAAGGCAAAAGAAAGTCAGAT 420
 OY 421 CTACTGGATGCTGATTTTCAACTATTAAGATTTGAATTAACATTTACAGAAACTGAG 480
 DB 421 CTACTGGATGCTGATTTTCAACTATTAAGATTTGAATTAACATTTACAGAAACTGAG 480
 OY 481 TTAGGAGAGATTTGCTGCTCATCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTG 540
 DB 481 TTAGGAGAGATTTGCTGCTCATCTTCTGATGCTGATGCTGATGCTGATGCTGATGCTG 540
 OY 541 AATTTCAATTTGATTCATCTCATGATTTATTTGAATTTGATTCATCTTCTGATGCTGAT 600
 DB 541 AATTTCAATTTGATTCATCTCATGATTTATTTGAATTTGATTCATCTTCTGATGCTGAT 600
 OY 601 ACCAGACGTTAATGTTCTGTAATAACTTAATTCCTCATATTTGAGAGAAAGAGAGTG 660
 DB 601 ACCAGACGTTAATGTTCTGTAATAACTTAATTCCTCATATTTGAGAGAAAGAGAGTG 660
 OY 661 ATGACGATTAATTTGAGCCCTTCTCTGCTGCTGCAAACTGTAAGGAGACAAAG 720
 DB 661 ATGACGATTAATTTGAGCCCTTCTCTGCTGCTGCAAACTGTAAGGAGACAAAG 720
 OY 721 GTGGCAGAAATTAACACGAGAGACTTAAGATGCTTTTACCGTTGATTCACCACTT 780
 DB 721 GTGGCAGAAATTAACACGAGAGACTTAAGATGCTTTTACCGTTGATTCACCACTT 780
 OY 781 TAACTAATGCGACAGAAATGGAATCACCTTAATTAATGATGAGAAACCGTGTTAGAT 840
 DB 781 TAACTAATGCGACAGAAATGGAATCACCTTAATTAATGATGAGAAACCGTGTTAGAT 840
 OY 841 ATCCGAGCAACAGAGCTTGTGACAAATTTGAGATGATGATGATGATGATGATGATGAT 900
 DB 841 ATCCGAGCAACAGAGCTTGTGACAAATTTGAGATGATGATGATGATGATGATGATGAT 900
 OY 901 TAGCCACTCTTCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 TAGCCACTCTTCTGAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 OY 961 CTGCTACTAAGCAAGAAAGTTATCTGCTGAGCCTGCTGAGCCTGCTGAGCCTGCTGAG 1020
 DB 961 CTGCTACTAAGCAAGAAAGTTATCTGCTGAGCCTGCTGAGCCTGCTGAGCCTGCTGAG 1020
 OY 1021 TTCAAGAAAGCCCAATTTATCATCTTCTACAGAGTCCAGCAATGGCGCTTGAGCAAGTAA 1080
 DB 1021 TTCAAGAAAGCCCAATTTATCATCTTCTACAGAGTCCAGCAATGGCGCTTGAGCAAGTAA 1080
 OY 1081 TGTCTGACCGGAGATGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1140
 DB 1081 TGTCTGACCGGAGATGAGGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATG 1140
 OY 1141 AGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 DB 1141 AGATGCTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 OY 1201 ACTGCTTGTAAAGAAACAAAGGCTTATAGCAAGATGATGATGATGATGATGATGATGAT 1260
 DB 1201 ACTGCTTGTAAAGAAACAAAGGCTTATAGCAAGATGATGATGATGATGATGATGATGAT 1260
 OY 1261 CATTTTCAAGATTTTACAGAAAGATGACACCGTTACTCATCTCTTCTGAGTGTGGA 1320
 DB 1261 CATTTTCAAGATTTTACAGAAAGATGACACCGTTACTCATCTCTTCTGAGTGTGGA 1320
 OY 1321 GATTTGTTTATTAATTAATGAGCAACATGATGATGATGATGATGATGATGATGATGATG 1380
 DB 1321 GATTTGTTTATTAATTAATGAGCAACATGATGATGATGATGATGATGATGATGATGATG 1380
 OY 1381 GCAATTAACATCTGAGAAATTTGCGGTAATAGCTGAGACACCAACCAACCAACCA 1440

```

|||||
Db 1381 GCAATACCATCTCGAGATTCGCCCTATAGCTGACACACACACCAACCAACAA 1440
OY 1441 ACAGTGTGATGTCGCCAGTACACCAACACCAACCAATATATGATATATCCCA 1500
Db 1441 ACAGTGTGATGTCGCCAGTACACCAACCAACCAATATATGATATATCCCA 1500
OY 1501 ATGACATTAACACCAAGAACATGTTGACCAACAGACATTAACAGACAGACAA 1560
Db 1501 ATGACATTAACACCAAGAACATGTTGACCAACAGACATTAACAGACAGACAA 1560
OY 1561 TTAATAGGAAATATCGGCTTTATGATACCGATTTATGATGATTTATCTTC 1620
Db 1561 TTAATAGGAAATATCGGCTTTATGATACCGATTTATGATGATTTATCTTC 1620
OY 1621 TTTCTTAAATAATGTTTATGAGACCAACAAATTTTATGATGATTTATCAACTGAT 1680
Db 1621 TTTCTTAAATAATGTTTATGAGACCAACAAATTTTATGATGATTTATCAACTGAT 1680
OY 1681 TACATTTTATGTTTAAATAATGATTTCTGCTTATTA 1722
Db 1681 TACATTTTATGTTTAAATAATGATTTCTGCTTATTA 1722

```

RESULT 3

AAA63671 ID AAA63671 standard; cDNA; 1715 BP.

AAA63671;

04-DEC-2000 (first entry)

Nucleotide sequence of VRN2 sequence of Arabidopsis sp.

H51; one locus-FRIGIDA; FRI gene; flowering time; blotting; flower initiation; stem elongation; flower production; VRN2; ss.

Arabidopsis sp.

WO20046358-A2.

10-AUG-2000.

25-JAN-2000; 2000WO-GB00197.

05-FEB-1999; 99GB-0002660.

(PLAN-) PLANT BIOSCIENCE LTD.

Johanson U, West J, Dean C;

WPI; 2000-532899/48.

New nucleic acid derived from the FRI locus of a plant, e.g. Arabidopsis, encoding a polypeptide capable of specifically altering the flowering time of a plant -

Example 5; Page 54-55; 73pp; English.

The present sequence represents a VRN2 sequence from variety Columbia. The specification describes a sequence which encodes a polypeptide capable of specifically altering the flowering time of a plant. The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus of Arabidopsis. The FRI polynucleotide is used to transform plants, so that the flowering time of a plant is altered. This is used, for example, for plants in which the leaves or tubers are a commercial product, where it is desirable to avoid 'blotting' (initiation of flowers and stem elongation) at too early a stage. Conversely, it may be desirable to alter flowering under certain circumstances e.g. to vary flower production across the seasons.

Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;

```

Query Match 91.1%; Score 1569.6; DB 21; Length 1715;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;
OY 1 CAAGCTTCTCAATTTTGGCTTG--CTCTCTTACACAGCCAAATCGGTGTTGCACT 58
Db 1 CAAGCTTCTCAATTTTGGCTTGCTCTCTCTTACACAGCCAAATCGGTGTTGCACT 60
OY 59 TTCAGGCTTCATCCAAAGACATTTCTATATATATATATATGACAAAGAGCGGTTCAATG 118
Db 61 TTCAGGCTTCATCCAAAGACATTTCTATATATATATATATGACAAAGAGCGGTTCAATG 120
OY 119 TTGCATGAGTTTATGCTATGATGAGGAAATTTCTATTTAGGGAGGCGTCAAGATT 178
Db 121 TTGCATGAGTTTATGATGAGGAAATTTCTATTTAGGGAGGCGTCAAGATT 180
OY 179 TTGCATGAGTTTATGCTATGATGAGGAAATTTCTATTTAGGGAGGCGTCAAGATT 238
Db 181 TTGCATGAGTTTATGCTATGATGAGGAAATTTCTATTTAGGGAGGCGTCAAGATT 240
OY 239 GCAGAAATGTCGGCGGCAAAATCCGACCCGAGAGAGATGATTTCACTGATGAAATCTCT 298
Db 241 GCAGAAATGTCGGCGGCAAAATCCGACCCGAGAGAGATGATTTCACTGATGAAATCTCT 300
OY 299 GATATATGTTAAACCTGTTGACATATATATATATATATATATATATATATATATAT 358
Db 301 GATATATGTTAAACCTGTTGACATATATATATATATATATATATATATATATATATAT 360
OY 359 ATCGTTTCTCCAAAGATGCTTGAATACAAATTTGAGCAAAACGCAAAAGAAATCAAG 418
Db 361 ATCGTTTCTCCAAAGATGCTTGAATACAAATTTGAGCAAAACGCAAAAGAAATCAAG 420
OY 419 ATCTACTGGAGATGATGATTTTCACTATATATATATATATATATATATATATATATAT 478
Db 421 ATCTACTGGAGATGATGATTTTCACTATATATATATATATATATATATATATATATAT 480
OY 479 AGTTAGGAGAGATGATGATTTTCACTATATATATATATATATATATATATATATATAT 538
Db 481 AGTTAGGAGAGATGATGATTTTCACTATATATATATATATATATATATATATATATAT 540
OY 539 GCAATTTCAATTTGAATATCATCATGATTTATATATATATATATATATATATATATAT 598
Db 541 GCAATTTCAATTTGAATATCATCATGATTTATATATATATATATATATATATATATAT 600
OY 599 ATACGACAGATTAATGTTTGTGTAATTAATTTCTTCAATTTTGGAGAAAGAGAG 658
Db 601 ATACGACAGATTAATGTTTGTGTAATTAATTTCTTCAATTTTGGAGAAAGAGAG 660
OY 659 TGATGACGATTAATTTGAGCCCTTCTCTGCTGCGAAGCCCTGTAAGCGAGACAAAG 718
Db 661 TGATGACGATTAATTTGAGCCCTTCTCTGCTGCGAAGCCCTGTAAGCGAGACAAAG 720
OY 719 AGGTGACAGAAATTAACACAGAGAGACTTAAGATATGCTTTTACCGTTGATTCACCCAG 778
Db 721 AGGTGACAGAAATTAACACAGAGAGACTTAAGATATGCTTTTACCGTTGATTCACCCAG 780
OY 779 TTTAATGATGACAGAAATTAAGATATGCTTTTACCGTTGATTCACCGTTGATTCAG 838
Db 781 TTTAATGATGACAGAAATTAAGATATGCTTTTACCGTTGATTCACCGTTGATTCAG 840
OY 839 ATATCCCGGAGCAAGAGCTGTCGCAATTTAGATATATATATATATATATATATATAT 898
Db 841 ATATCCCGGAGCAAGAGCTGTCGCAATTTAGATATATATATATATATATATATATAT 900
OY 899 CATAGCCACTTCTCTGAGCGTGTGCTAAATTTATATATATATATATATATATATATAT 958
Db 901 CATAGCCACTTCTCTGAGCGTGTGCTAAATTTATATATATATATATATATATATATAT 960
OY 959 CCTGCTACTAGACAGAAAGATTTATCTGCTGAGCGATCAGAGGCTTAAAGCCACTACT 1018
Db 961 CCTGCTACTAGACAGAAAGATTTATCTGCTGAGCGATCAGAGGCTTAAAGCCACTACT 1020
OY 1019 TCTTGAAGAGCCAAATTTCTATCTATCTACAGAGATCCAGCAATGGCGTTGAGCAAGT 1078

```

[illegible]

XX	Dean C,	Gendall A;
PI		
DR	WPI; 2000-499333/4.	
DR	P-FSDB; AAB00061.	
XX		
PT	Isolated vernalization gene VRN2 is used to produce transgenic plants	
PT	with altered vernalization response, flowering time, leaf size and/or	
PT	shape or shade avoidance response for maximized reproductive success	
XX		
PS	Claim 5; Page 71; 105pp; English.	
XX		
CC	Isolated nucleic acid sequences obtained from the VRN2 locus of a	
CC	plant encode polypeptides which are capable of affecting one or more	
CC	vernalization responses such as, flowering time, leaf size and/or	
CC	shape or the shade avoidance response of a plant into which the	
CC	nucleic acid is introduced. Introducing such sequences into plants	
CC	to alter these characteristics maximises the reproductive success of	
CC	the plant.	
XX		
SQ	Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other:	
Query Match	91.1%; Score 1569.6; DB 21; Length 1715;	
Best Local Similarity	96.5%; Pred. No. 0;	
Matches 1649; Conservative	0; Mismatches 49; Indels 10; Gaps 4.	
OY	1 CAAGCTTCTTCAAATTTGCTGG--CTCCTCTTTACACGCCCAATGGGTTCGCAGCT	58
Dd		
Dd	1 CAAGCTTCTTCAAATTTGCTTGCTCTCTCTCTTTACAGGCCCAATGGGTTCGCAGCT	60
OY	59 TTCAGGCCCTCAATCCAACACATCTCATTAAGCATATTGCACAGAAGCGGTTCTAATTG	118
Dd	61 TTCAGGCCCTCAATACACACATCTCATTAAGCATATTGCACAGAAGCGGTTCTAATTG	120
OY	119 TTGCATTAGTTTATCGCTATGACGTAGGGAATTTCTAATTTAGGGAGGCTTCAGAGTT	178
Dd	121 TTGCATVAGAGTGAACAATATGACGTACGAGGAATTTCTAATTTAGGGAGGCTTCAGAGTT	180
OY	179 TGCACTAATCTCATATATGCGCTCTGACGTTGTGAGTGTATAATGAAACAGATGTGTAG	238
Dd	181 TGCACTAATCTCATATATGACGCTCTGACGTTGTGAGTGTATAATGAAACAGATGTGTAG	240
OY	239 GCAGCAATTTGCGCGGAAATCTCTCACCGGAGAGTAGATTTCACACTGATGAGATCTCTT	298
Dd	241 GCAGCAATTTGCGCGGAAATCTCTCACCGGAGAGTAGATTTCACACTGATGAGATCTCTT	300
OY	299 GATATATTGTAAACCTGTTGCACTGTATTAACATCTTACCCTTGCTCTTAGGCAACC	358
Dd	301 GATATATTGTAAACCTGTTGCACTGTATTAACATCTTACCCTTGCTCTTAGGCAACC	360
OY	359 ATCGTTTTCTCCAAAGATCTTGAACATACAAAATTTGGACCAAAGCCAAAGAAAGTCAAG	418
Dd	361 ATCGTTTTCTCCAAAGATCTTGAACATACAAAATTTGGGCAAAAGCCAAAGAAAGTCAAG	420
OY	419 ATCTACTGTGGAGTGTAGTTTTCACATTAAGAGTGTATTAACATTTACAGAAAACTGA	478
Dd	421 ATCTACTGTGGAGTGTAGTTTTCACATTAAGAGTGTATTAATPATATACAAAGAACTGA	480
OY	479 AGTTAGGAGAGATTGTTCTTGCTCAATTTTGGCTATGCTATGTTAGTCTCAAGGGCT	538
Dd	481 AGTTAGGAGAGATTGTTCTTGCTCAATTTTGGCTATGCTATGTTAGTCTCAAGGGCT	540
OY	539 GCAATTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGITCAAGCTTTTGAAGA	598
Dd	541 GCAATTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGITCAAGCTTTTGAAGA	600
OY	599 ATACACAGACATTAATGTTTCTGTAATAAATCTTAATCTTCAATTTTGGAGAAAGAGAG	658
Dd	601 ATACACAGACATTAATGTTTCTGTAATAAATCTTAATCTTCAATTTTGGAGAAAGAGAG	660
OY	659 TGATGACGATAAATTTGAGCCCTTCTCTCTGCTCGAAGCCCTGTAAGCGGAGCAAG	718
Dd	661 TGATGATATPAAATTTGAGCCCTTCTCTCTGCTCGAAGCCCTGTAAGCGGAGCAAG	720

OY	719	AGGTGGCAGAAATATACACAGAGACTAAAGTATGCTTTTACGTTGGATTCCACCAG	778
Db	721	AGGTGGCAGAAATAATACACAGAGACTTAAAGTATGCTTTTACGTTGGATTCCACCAG	780
OY	779	TTTAACTATATGGCAGAGAAATGSAATACCCCTACTTAATGATGTAAGAACCGGGTTTAG	838
Db	781	TTTAACTATATGGCAGAGAAATGSAATGATTCCTCGTGAATGATGTAAGAACCGGGTTTAG	840
OY	839	ATATCCGAGGCAACAGAGCTTGTGGACAAATTTAGATGACACCAATTCACACAG	898
Db	841	ATATCCGAGGCAACAGAGCTTGTGGACAAATTTAGATGATGATGTAAGAACCGGGTTTAG	900
OY	899	CATAGCCACACTCTCTCTGGACGCTGTGCTTAAAGTTATATTTAGACAAGGAAAGTGTGT	958
Db	901	CATAGCCACACTCTCTCTGGACGCTGTGCTTAAAGTTATATTTAGACAAGGAAAGTGTGT	960
OY	959	CCCTGCTACTAAGACAAAGAAATTAATCTGCTGAGGATCCAGAGGCTAGAAAGCCACTCT	1018
Db	961	CCCTGCTACTAAGACAAAGAAATTAATCTGCTGAGGATCCAGAGGCTAGAAAGCCACTCT	1020
OY	1019	TCTTCAGAAAGCCCAATTCATCTCACAGATCCAGCAATCCAGCAATGGCTTGACACAGT	1078
Db	1021	TCTTCAGAAAGCCCAATTCATCTCACAGATCTCACAGATCCAGCAATGGCTTGACACAGT	1080
OY	1079	AATGCTACCCGGGGATACGAGAGATGAGTGCATGATGATGATGATGATGATGATGATG	1138
Db	1081	AATGCTACCCGGGGATACGAGAGATGAGTGCATGATGATGATGATGATGATGATGATG	1140
OY	1139	CCAGATGCTTATGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG	1198
Db	1141	CCAGATGCTTATGATGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATG	1200
OY	1199	GACCTCGTTTATGAGAAACAAAGGTTATGACAAATGCTCATCTCTGGGCAATGGA	1258
Db	1201	GACCTCGTTTATGAGAAACAAAGGTTATGACAAATGCTCATCTCTGGGCAATGGA	1260
OY	1259	AGCAATTCAGATTTTACGAGAAAGATGACCGTTACTCATCTCTGGTGTG	1318
Db	1261	AGCAATTCAGATTTTACGAGAAAGATGACCGTTACTCATCTCTGGTGTG	1320
OY	1319	GAGATTTGTTTATTAATTAATGAGAACCATGAGCTTTCGACTGAGCCACCATCAACA	1378
Db	1321	GAGATTTGTTTATTAATTAATGAGAACCATGAGCTTTCGACTGAGCCACCATCAACA	1380
OY	1379	CTGCAATACCATCTCGAAGATTTGCCGTAATAGCTCAGACACACACACACACACAA	1438
Db	1381	CTGCAATACCATCTCGAAGATTTGCCGTAATAGCTCAGACACACACACACACACAA	1440
OY	1439	CAACAGTGTGATGCTCCAGCTGACTCAACACCAACCAATATACATTTGATGATCAATC	1498
Db	1441	CAACAGTGTGATGCTCCAGCTGACTCAACACCAACCAATATACATTTGATGATCAATC	1500
OY	1499	CAATGACATTAACAACAAGACAAATGTTGACACAAAGACATATACAGAGACAAAGT	1558
Db	1501	CAATGACATTAACAACAAGACAAATGTTGACACAAAGACATATACAGAGACAAAGT	1560
OY	1559	AATTAATATAGAAATCTCCGGCTTTTATGATACCGATTAATGATGATGATGATGATG	1618
Db	1561	AATTAATATAGAAATCTCCGGCTTTTATGATACCGATTAATGATGATGATGATGATG	1620
OY	1619	TCTTTCTT-AAAAATGTTTAAAGACCAACAATTTTATATATGTTAGTATCAACT	1678
Db	1621	TCTTTCTT-AAAAATGTTTAAAGACCAACAAGATTTTATATGTTAGTATCAACT	1680
OY	1679	GATTACATTTTAAAGAAAAAAT 1705	
Db	1681	GATTACATTTTAAAGAAAAAAT 1701	

RESULT 5
AAA47755
ID AAA47755 standard; cDNA; 1737 BP

XX	AAAA7755:
AC	16-NOV-2000 (first entry)
DT	
XX	Aberrantly spliced VRN2 nucleic acid.
DE	
XX	Vernalization gene: VRN2; plant characteristic: flowering time;
KW	leaf size; leaf shape; shade avoidance response; reproduction;
KW	breeding; pollination; cultivation; ss.
XX	
OS	Arabidopsis thaliana var 'Columbia'.
XX	MO200044918-A1.
FN	
PD	03-AUG-2000.
XX	
PF	28-JAN-2000; 2000WO-GB00248.
XX	
PR	28-JAN-1999; 99GB-0001927.
XX	
PA	(PLAN-) PLANT BIOSCIENCE LND.
XX	
PI	Dean C, Gendall A;
XX	
DR	WPL: 2000-499333/44.
DR	P-PDB: AAB00062.
XX	
PT	Isolated vernalization gene VRN2 is used to produce transgenic plants
PT	with altered vernalization response, flowering time, leaf size and/or
PT	shape or shade avoidance response for maximized reproductive success
XX	
FS	Claim 8: Page 74-75; 105pp: English.
XX	
CC	Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC	plant encode polypeptides which are capable of affecting one or more
CC	vernmalization responses such as, flowering time, leaf size and/or
CC	shape or the shade avoidance response of a plant into which the
CC	nucleic acid is introduced. Introducing such sequences into plants
CC	to alter these characteristics maximises the reproductive success of
CC	the plant. This cDNA was generated from an aberrantly spliced VRN2
CC	gene.
XX	
SQ	Sequence 1737 BP; 543 A; 330 C; 359 G; 505 T; 0 other;
	Query Match 89.3%; Score 1537.6; DB 21; Length 1737;
	Best Local Similarity 95.3%; Pred. NO. 0;
	Matches 1649; Conservative 0; Mismatches 49; Indels 32; Gaps 5
OY	1 CAAGCTTCTCAATTTTGCTG--CTGTCTTTACAGGCATCGGTGTTTCGAGCT 58
Db	1 CAAAGCTCTTCAATTTTGCTGTCTCTCTCTTACAGGCCAATCGGTGTTTCGAGCT 60
OY	59 TTCAGGCTCAATCCAGAACAATTCTATATAAGCATATTCAGAAAGGGGGTTCTAATG 118
Db	61 TTCAGGCCCAATACAAAGACATCTATATAAGCATATTCAGAAAGGGGGTTCTAATG 120
OY	119 TTGCATGTGAATTAATCGGTATAGAGTAGGAAAATCTAATTTAGGGAGAGCCTCAGAGTT 178
Db	121 TTGCATGTGAATTAATCGGTATAGAGTAGGAAAATCTAATTTAGGGAGAGCCTCAGAGTT 180
OY	179 TCACACTAATCTATATATGCGCTCTTGAGAGTTGTTGAGCTGTAATTTGAACAAGATGTGTA 238
Db	181 TGACACTAATCTATATATGCGCTCTTGAGAGTTGTTGAGCTGTAATTTGAACAAGATGTGTA 240
OY	239 GCAGAAATGTGCGGGAATTCCTACCGGAGAGAGATTTGAACGATGAGAAATCTCT 298
Db	241 GCAGAAATGTGCGGGAATTCCTACCGGAGAGAGATTTGAACGATGAGAAATCTCT 300
OY	299 GATATATGTGAACCTGTTGAGATATATAACATCTTACCTCGCTCTAGGCAACC 358
Db	301 GATATATGTGAACCTGTTGAGATATATAACATCTTACCTCGCTCTAGGCAACC 360

OY	359	ATCGTTTCTCCAGATGCTTGAACTACAAAAATTGGACAAAGCCCAAGAAAGTCAAG	418
OY	361	ATCGTTTCTGCCAAGATGCTTGAAGCTACAAAAATTGGGGCAAAAGGCCAAAGAAAGTCAAG	420
OY	419	ATACACTGGAGGGAGTGTTCACACTAATAAGATGTGTAATTAACATTAACAGAAACGA	478
OY	421	ATCTACTGGATGTGATGTTCACTAACTAATAAGATGTGTAATTAACATTAACAGAAACGA	480
OY	479	AGTTAGGAGAGATGTTCTTGCCATATTTGGCTATGCTATGATGGTAGCTTCA-----	531
DB	481	AGTTAGGAGAGATGTTCTTGCCATATTTGGCTATGCTATGATGGTAGCTTCAAGGTGG	540
OY	532	-----AGGGGCTGCAATTTCAATTTGAATTCATCTCAGATTTAATTGAAT	576
DB	541	CAAGTATATACACTGAGGGGCTGGAATTTCAATTGAATTCATCTCAGATTTAATTGAAT	600
OY	577	TTGAGTTAACCTTTTGGAAATAACCGAGATTAATGTTCTGTAAACCTTAATTCCT	636
DB	601	TTGAGTTAACCTTTTGGAAATAACCGAGATTAATGTTCTGTAAACCTTAATTCCT	660
OY	637	TCATATTTGAGAGAAGAAAGAGTATGACGATTAATTTGAGCCCTTCTCTCTGCTGCA	696
DB	661	TCATATTTGAGAGAAGAAAGAGTATGATGATTAATTTGAGCCCTTCTCTCTGCTGCA	720
OY	697	AACTCGTAAGCGAGACAAAGAGGTGGCAGAAATACACAGAGAGACTTAAAGTATGCT	756
DB	721	AACTCGTAAGCGAGACAAAGAGGTGGCAGAAATACACAGAGAGACTTAAAGTATGCT	780
OY	757	TTTAAACGTTGGATTCACCCAGTTTAACATAATGGCACAGAAATAGGAATCAACCTACTA	816
DB	781	TTTAAACGTTGGATTCACCCAGTTTAACATAATGGCACAGAAATAGGAATTCACCTACTA	840
OY	817	ATGATGGAACCCGTGTTTAGATATCCCGAGGCAACAGACTTGCTGTGACAAATTTGGA	876
DB	841	ATGATGGAACCCGTGTTTAGATATCCCGAGGCAACAGACTTGCTGTGACAAATTTGGA	900
OY	877	TGACCAGCAATTCACACGACCAGTAAAGCCACTTCTCTGAGCGTGTGCTAAAGTTA	936
DB	901	TGACTAGCAATTCACACGACCAGTAAAGCCACTTCTCTGAGCGTGTGCTAAAGTTA	960
OY	937	TATTGACAAGGAAGCTGTGCTGCTGCTACTAAGACAAAGAAATATCTGCTGAGGAT	996
DB	961	TATTGACAAGGAAGCTGTGCTGCTGCTACTAAGACAAAGAAATATCTGCTGAGGAT	1020
OY	997	CAGAGGCTAGAGGACCACTTCTGTTGCAAAAGCCAAATTCATCATCTGACAGAGTCC	1056
DB	1021	CAGAGGCTAGAGGACCACTTCTGTTGCAAAAGCCAAATTCATCATCTGACAGAGTCC	1080
OY	1057	AGCCAAATGCGCTTGAGCAAGTAATGTCTGACCGGGATAGGAGAGATGAAGTCAGTACG	1116
DB	1081	AGCCAAATGCGCTTGAGCAAGTAATGTCTGATCGGGATAGGAGAGATGAAGTCAGTACG	1140
OY	1117	ATGTTGCAGATTTTGAAGATGCCAGATGCTTGATGACTTGTGGAATGTAATTAAGATG	1176
DB	1141	ATGTTGCAGATTTTGAAGATGCCAGATGCTTGATGACTTGTGGAATGTAATTAAGATG	1200
OY	1177	AAAAGCAATTCATCATCTTTGGAAACGCTTGTGAAGAAAACAAAGGTTTATACAGATG	1236
DB	1201	AAAAGCAATTCATCATCTTTGGAAACGCTTGTGAAGAAAACAAAGGTTTATACAGATG	1260
OY	1237	GTCATATCTCTTGGGCAATGTGAACATTTTCAGATTTTACGAGAAGAGTTGACCGT	1296
DB	1261	GTCATATCTCTTGGGCAATGTGAACATTTTTCAGATTTTACGAGAAGAGTTGACCGT	1320
OY	1297	ACTCATCACTTCTCTGGTGTGGAGATTTGTTTGATTAACCTATGCAACATGACATTTG	1356
DB	1321	ACTCATCACTTCTCTGGTGTGGAGATTTGTTTGATTAACCTATGCAACATGACATTTG	1380
OY	1357	TCGACTAGCCACCATCAACAACTGCAATACCATCTCGAGAATTTGGCGTAAATAGCTACG	1416
DB	1381	TCGACTAGCCACCATCAACAACTGCAATACCATCTCTGAGAAATTTGCCGTAAATACCTCAG	1440
OY	1417	ACACCAACCAACCAACAAACAAAGTGTGATGCTGCCAGTACTCAAAACCAACCA	1476

[illegible]

PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139453.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142390.
PR 09-JUL-1999; 99US-0142803.
PR 12-JUL-1999; 99US-0142820.
PR 13-JUL-1999; 99US-0142877.
PR 14-JUL-1999; 99US-0143342.
PR 15-JUL-1999; 99US-0143624.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144085.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 02-AUG-1999; 99US-0146386.
PR 99US-0146388.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149375.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 27-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 15-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154079.
PR 22-SEP-1999; 99US-0154139.
PR 23-SEP-1999; 99US-0154486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 21-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.

Fri Jun 20 09:03:38 2003

us-09-890-220-1.1mg

PR 28-OCT-1999: 9905-0161992.
 PR 28-OCT-1999: 9905-0161993.
 PR 29-OCT-1999: 9905-0162142.

56.9%; Score 979.8; DB 21; Length 1497;

Query Match Best Local Similarity 89.7%; Pred. No. 1.3e-232; Indels 102; Gaps 2;
 Matches 1118; Conservative 0; Mismatches 27;

355 ACCCATCGTTCTTCCAGATGCTTGAACCTACAAATTTGAGCAAGCCCAAAAGAAAGT 414
 DB 320 AGCCATCGTTCTTCCAGATGCTTGAACCTACAAATTTGAGCAAGCCCAAAAGAAAGT 379
 OY 415 CAAAGCTCTACTGGATGATGATGTTTCAACTATTAAGATTTAATACATTAACATTAACAGAAA 474
 DB 380 CAAATCTACTGGATGATGATGTTTCAACTATTAAGATTTAATACATTAACATTAACAGAAA 439
 OY 475 CTGAAGTTAGGAGATGTTCTTCTGCAATTTTGTCTATGCTATGCTATGCTTCAAG 534
 DB 440 CTGAAGTTAGGAGATGTTCTTCTGCAATTTTGTCTATGCTATGCTATGCTTCAAG 499
 OY 535 GGCATCATTTTCAATTTGATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTGATTTG 594
 DB 500 GGCATCATTTTCAATTTGATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTGATTTG 559
 OY 595 AAGATACCAAGCAATTTATGTTCTGTAATACTTAATTCCTTCATATTTGAGAAAG 654
 DB 560 CTGTAAATCCAAACATTA----- 577
 OY 655 GAAGTATGATGATTAATTTGAGCCCTTCTCTCTGCTGCAAACTCGTAAGCGGAGAC 714
 DB 578 -----1GTTTCCAGCTGCAAACTCGTAAGCGGAGAC 610
 OY 715 AAAGAGTGGCAGAAATTAACACCGAGAGACCTTAAGATGCTTTTACCGTGGATTCAC 774
 DB 611 AAAGAGTGGCAGAAATTAACACCGAGAGACCTTAAGATGCTTTTACCGTGGATTCAC 670
 OY 775 CCAGTTTAACTAATGGCAGAAATTAAGATGCTTTTACCGTGGATTCAC 834
 DB 671 CCAGTTTAACTAATGGCAGAAATTAAGATGCTTTTACCGTGGATTCAC 730
 OY 835 TAGATATCCGAGGAGCAAGAGCTTGTGACCAATTTGATTTGATTTGATTTGATTTGATTTG 894
 DB 731 TAGATATCCGAGGAGCAAGAGCTTGTGACCAATTTGATTTGATTTGATTTGATTTGATTTG 790
 OY 895 CAGCCATAGCCCACTCTCTCTGAGCGTGTCTAAAGTATTTGACAAAGGAGAGCTG 954
 DB 791 CAGCCATAGCCCACTCTCTCTGAGCGTGTCTAAAGTATTTGACAAAGGAGAGCTG 850
 OY 955 TGGTCCCTGCTACTAAGACAAAGATTTATGCTGAGCGTGTCTAAAGTATTTGACAAAGGAG 1007
 DB 851 TGGTCCCTGCTACTAAGACAAAGATTTATGCTGAGCGTGTCTAAAGTATTTGACAAAGGAG 910
 OY 1008 -----AGCCACCTCTCTCTGAGCGTGTCTAAAGTATTTGACAAAGGAGAGCTG 1041
 DB 911 CCATATTAACGAGAAATTTCTTCCAGCCACCTTCTCTGAGCGTGTCTAAAGTATTTGACAAAG 970
 OY 1042 ATTCTCAGAGAGTCCAGCAATGCGCTTGAAGCAATATGCTTCAACCGGGTGAAGCGAG 1101
 DB 971 ATTCTCAGAGAGTCCAGCAATGCGCTTGAAGCAATATGCTTGAAGCAATATGCTTGAAGCA 1030
 OY 1102 ATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1161
 DB 1031 ATGAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1090
 OY 1162 ATGTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1221
 DB 1091 ATGTGAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1150
 OY 1222 GGGTTATAGCAATGCTATCTCTTGGGCAATGTAAGCAATTTTCAAGTTTACAGAG 1281
 DB 1151 GGGTTATAGCAATGCTATCTCTTGGGCAATGTAAGCAATTTTCAAGTTTACAGAG 1210
 OY 1282 AAGAGTTGACACCGTACTCATCTCTCTGCTGTTGGAGATGTTTGTGATTAACAT 1341

DB 1211 AAGATTCCACCGTTACTCATCTCTTGTGGTGGAGATGTTTGTGATTAACAT 1270
 OY 1342 GGAACCATGAGCTTGTGACTACGACCATCAATCAACATGCAATCACTCGAGAA 1401
 DB 1271 GGAACCATGAGCTTGTGACTACGACCATCAATCAACATGCAATCACTCGAGAA 1330
 OY 1402 GCGTATATGCTCAGACACGACGACGACGACGACGACGACGACGACGACGACG 1461
 DB 1331 GCGTATATGCTCAGACACGACGACGACGACGACGACGACGACGACGACG 1390
 OY 1462 ACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1521
 DB 1391 ACTCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1450
 OY 1522 ATGTTGACAAAGCAATTAACATTAACATTAACATTAACATTAACATTAAC 1568
 DB 1451 ATGTTGACAAAGCAATTAACATTAACATTAACATTAACATTAACATTAAC 1497

RESULT 7
 ID AAA47752 standard; DNA; 6338 BP.
 AC AAA47752;
 XX 16-NOV-2000 (first entry)
 DE VRN2 nucleic acid.
 XX VRN2 nucleic acid.
 DE VRN2 nucleic acid.
 XX VRN2 nucleic acid.
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KM breeding; pollination; cultivation; ds.
 OS Arabidopsis thaliana var 'landsberg erecta'.
 PN WO200044918-A1.
 PD 03-AUG-2000.
 PF 28-JAN-2000; 2000MO-GB00248.
 PR 28-JAN-1999; 99GB-0001927.
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 PI Dean C, Gendall A;
 PS WPI: 2000-499333/44.
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 PS Claim 9; Page 68-70; 105pp; English.
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as flowering time, leaf size and/or
 CC shape or shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.
 SQ Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;

Query Match 23.9%; Score 412.2; DB 21; Length 6338;
 Best Local Similarity 98.8%; Pred. No. 9.4e-92;
 Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0,
 OY 1304 ACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1363
 DB 5102 AATTTCCAGGTTGGAGATGTTTGTGATTAACATTAAGCAACATGACTGTGACTC 5161

Query Match 18.5%, Score 319.4; DB 21; Length 5895;

	Best Local Similarity	92.7%	Fred. No. 8.9e-69;	
Matches 370;	Conservative 0;	Mismatches 21;	Indels 8;	Gaps 3;
QY	1308	TTCTGTGTGTGGAGATTTGTTTGTATTAACATAATGGACCATGGACTTTCGACTACGC	1367	
Db	4677	TCGAGTGTGGAGATTTGTTTGTATTAACATAATGGACCATGGACTTTCGACTACGC	1367	
QY	1368	ACCAATCAACACACACATACCATTCCTCGAATTTGCCATTAATAGCTCAGACACACAC	1427	
Db	4737	ACCAATCAACACACATACCATTCCTCGAATTTGCCATTAATAGCTCAGACACACAC	1427	
QY	1428	ACCAATCAACACACATACCATTCCTCGAATTTGCCATTAATAGCTCAGACACACAC	1427	
Db	4791	ACCAATCAACACACATACCATTCCTCGAATTTGCCATTAATAGCTCAGACACACAC	1427	
QY	1488	GTGATCATTCGCCAATGACATTAACAAACAAACAAACATTTGACATTAACATTA	1487	
Db	4851	GTGATCATTCGCCAATGACATTAACAAACAAACAAACATTTGACATTAACATTA	1487	
QY	1548	AGAGACAAGTAATTAATAGAAAATCTCCGCTTTTATGATACGATTTATGCGATG	1607	
Db	4911	AGAGACAAGTAATTAATAGAAAATCTCCGCTTTTATGATACGATTTATGCGATG	1607	
QY	1608	TACTTATTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1666	
Db	4970	TACTTATTTCTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	1666	
QY	1667	TGATTCACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1705	
Db	5030	TGATTCACATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA	1705	
RESULT 9				
AAC51582				
ID	AAC51582	standard; DNA; 319 BP.		
AC	AAC51582:			
XX	18-OCT-2000	(first entry)		
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 69042.			
XX	Hybridisation assay; genetic mapping; gene expression control;			
KW	protein identification; signal transduction pathway;			
XX	metabolic pathway; promoter; termination sequence; ss.			
OS	Arabidopsis thaliana.			
XX	EP1033405-A2.			
PN	06-SEP-2000.			
XX	25-FEB-2000; 2000EP-0301439.			
XX	25-FEB-1999; 99US-0121825.			
PR	05-MAR-1999; 99US-0123180.			
PR	09-MAR-1999; 99US-0123548.			
PR	22-MAR-1999; 99US-0125788.			
PR	23-MAR-1999; 99US-0126264.			
PR	01-APR-1999; 99US-0126785.			
PR	08-APR-1999; 99US-0127462.			
PR	08-APR-1999; 99US-0128234.			
PR	19-APR-1999; 99US-0128714.			
PR	21-APR-1999; 99US-0129845.			
PR	23-APR-1999; 99US-0130077.			
PR	23-APR-1999; 99US-0130449.			
PR	28-APR-1999; 99US-0130891.			
PR	30-APR-1999; 99US-0131449.			
PR	04-MAY-1999; 99US-0132048.			
PR	05-MAY-1999; 99US-0132407.			
PR	05-MAY-1999; 99US-0132484.			
PR	05-MAY-1999; 99US-01332485.			

PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134270.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137502.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.

PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 18.2%; Score 314.2; DB 21; Length 319;
Best Local Similarity 99.1%; Pred. No. 5.4e-68;
Matches 316; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1375 ACAACGCAATACCATCTCGAGATTGCGTAATAGCTCAGACACGACGACGACACA 1434
DB 1 ACAACTGCAATACCATCTCGAGATTGCGTAATAGCTCAGACACGACGACGACACA 60
OY 1435 ACAACAGAGTGTGATCGTCGAGTACTCAACACGACGACGACGACGACGACGAC 1494
DB 61 ACAACAGAGTGTGATCGTCGAGTACTCAACACGACGACGACGACGACGACGACGAC 120
OY 1495 ATCCGATGATTAAC 1554
DB 121 ATCCGATGATTAAC 180
OY 1555 AAGTAATTAATAGGAAATCTCGGCTTTATGATACCGATTATCGATTGTACTTA 1614
DB 181 AAGTAATTAATAGGAAATCTCGGCTTTATGATACCGATTATCGATTGTACTTA 240
OY 1615 TTCTCTCTCTTAAATAATGTTAGACAAACAATTTTATATGTATGTTAGTATCA 1674
DB 241 TTCTCTCTCTTAAATAATGTTAGACAAACAATTTTATATGTATGTTAGTATCA 300
OY 1675 ACTGATTACATTTTACTT 1693
DB 301 ACTGATTACATTTTACTT 319

RESULT 10

AAC44147
ID AAC44147 standard; DNA: 640 BP.

AC AAC44147;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41806.

XX Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130591.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 01-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140961.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.

```
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144336.
PR 20-JUL-1999; 99US-0144337.
PR 20-JUL-1999; 99US-0144338.
PR 20-JUL-1999; 99US-0144339.
PR 21-JUL-1999; 99US-0144340.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145090.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145291.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146387.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149928.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0156458.
PR 28-SEP-1999; 99US-0156559.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
```

```
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 17.8%; Score 306.4; DB 21; Length 640;
Best Local Similarity 98.8%; Pred. No. 6.1e-66;
Matches 318; Conservative 1; Mismatches 2; Indels 1; Gaps 1;

QY 1373 CACACATGCAATACCATCTCGAGATTCGCGTAATAGCTCAGACACACACACACAA 1432
    |||||||
DB 319 CACACATGCAATACCATCTCGAGATTCGCGTAATAGCTCAGACACACACACACAA 378
    |||||||
QY 1433 CACACACACAGTGTGCTGCCAGTCAACACCAACATATACATTGTGGA 1492
    |||||||
DB 379 CACACACACAGTGTGCTGCCAGTCAACACCAACATATACATTGTGGA 438
    |||||||
QY 1493 TCATCCCATGACATTAACAACAAGACATGTTGACACACAGACATTAACAGACAG- 1551
    |||||||
DB 439 TCATCCCATGACATTAACAACAAGACATGTTGACACACAGACATTAACAGACAG- 498
    |||||||
QY 1552 ACAAGATTAATTAAGAAATCTCCGCTTTATGATACGATTTATCGATTGTAAC 1611
    |||||||
DB 499 ACAAGATTAATTAAGAAATCTCCGCTTTATGATACGATTTATCGATTGTAAC 558
    |||||||
QY 1612 TTATTCCTCTCTTAATAAATGTTTAGAGCAACAAATTTTATATGTTAGTGAT 1671
    |||||||
DB 559 TTATTCCTCTCTTAATAAATGTTTAGAGCAACAAATTTTATATGTTAGTGAT 618
    |||||||
QY 1672 TCAACTGATTACATTTTACTT 1693
    |||||||
DB 619 TCAACTGATTACATTTTACTT 640
    |||||||

RESULT 11
AAZ36947
ID AAZ36947 standard; cDNA: 2280 BP.
XX
AC AAZ36947;
XX
DT 13-MAR-2000 (first entry)
XX
DE cDNA encoding a MPCL protein having flowering regulating activity.
XX
KW Flowering regulating activity; MPCL; flowering; germination;
KW super early flowering mutation; altered flowering time;
KW flowering regulating gene; food crop; vegetable; flowering inhibition;
KW productivity; ss.
XX
```

us-09-890-220-1.rng

Page 14

[illegible]

PR 24-JUN-1999; 99JP-0179043.
 XX (MITA) MITSUI CHEM INC.
 XX
 XX Yoshiida N, Kato Y, Takahashi S, Yanai Y, Hirasuka J, Miwa T;
 DR WPI; 2000-064612/06.
 DR P-PSDB; AAV53933.
 XX
 PT Novel DNA used to produce transgenic plants with altered floral
 PI regulation which can have increased crop yields -
 XX
 PS Claim 6; Page 40-44; 53pp; English.
 CC The present sequence encodes a protein having a flowering regulating
 CC activity, which is designated Os-MPCI. An Arabidopsis MPCI is also
 CC disclosed in the specification. The rice and Arabidopsis cDNAs show
 CC significant homology with each other. A naturally occurring mutation
 CC of the MPCI gene eliminates normal flowering regulating ability of
 CC plants, and leads to flowering immediately after germination (super
 CC early flowering mutation). The MPCI polynucleotide sequence can be
 CC used to produce plants with altered flowering times in comparison
 CC with wild type plants, by enhancing or inhibiting the expression of
 CC the flowering regulating gene. Antisense polynucleotides can be used
 CC to reproduce the effects of the mutated MPCI gene. This alteration
 CC can be used to increase the yield of food crops. Flowering inhibition
 CC of vegetables increases their productivity.
 XX
 SO Sequence 2248 BP; 641 A; 454 C; 493 G; 660 T; 0 other;

Query Match 14.7%; Score 252.4; DB 21; Length 2248;
 Best Local Similarity 56.2%; Pred. No. 2.4e-52;
 Matches 555; Conservative 0; Mismatches 376; Indels 57; Gaps 2;

QY 424 CTGGGATGTTGTTTCACTATTAAGATTGATTAACACATTACAGAAACTGAAGTTA 483
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 939 CTGGCAATGCTTTTAACTACAAATACAAATACAAATACAAATACAAATACAAATAC 998
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 484 GGGAGGATGTTGTTGCTGCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 543
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 999 CTGAAGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1058
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 544 TTCAATTAATTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 603
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1059 GTCACTTAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1118
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 604 AGACATTAATGTTCTGTAATACTTAATTCCTCATATTTGAG--GAAGAAGAGAGTG 660
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1119 AGGCTGTAAATGTTAGTCTGAAGACTGATTTCTTGAGAACAGAGCTTTTGGCTGAGGAG 1178
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 661 ATGAGATTAATTTGAGCCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1179 TTGATCCAGACATTAACATTTTGTACCGCTCAAGATTTAAGAGGGTAAAGGGTGG 1238
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 721 GTGGCAGAAATTAACACGAGACTTAAAGTATGCTTTTACCTTTGATTCACCCAGTT 780
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1239 AAATCTCAAGATTAATTAAGCATGTACATCCACATATGTGATTCAGGATACCTG 1298
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 781 TAACATATGACAGAAAATGATCACCTACTTAATGATGAGAAACCGTGTAGAT 840
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1299 AAGATGCCAGCAGAGATCTGAAAGCATTTACGTGCAAGAGGAAATATGTTGTTGAG 1358
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 841 ATCCGAGAGCAAGAGCTGTGCGCAATTTGAGATGACCAAGCAAACTCCACAGCA 900
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1359 CACACGCTTCTGTGATCTCTGTAATTCATTAACAGGTAGCAATCTTTCAGAAC-- 1413
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 901 TAGCCCACTTCTCTGACGCTGCTCTAAAGTAAATTAATTAATTAATTAATTAATTA 960
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1414 -----AACAGTTTAC 1424
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 961 CTGCTACTAAGCAAGAAATTAATCTGCTGACGATCAAGAGCTGAAGCCACTACTTC 1020
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1425 AGTTGGAGACAGAAAGCTGCTGTTGAACGAGCTGATCCAGAAATGCGGACCTCC 1484
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

QY 1021 TTCAGAAAGCCCAATTCATTCATTCACAGAGTCACGCAATGCGGCTTGAGCAAGTAA 1080
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1485 TACAAAAGCCAGTCTTCTTCTACAGAGGCTCCACCAATGGCAGAGCAAGTTT 1544
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1081 TGTCTGACCGGATAGCAGAGATGAATGATGATGATGATGATGATGATGATGATGATG 1140
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1545 TCTCAGATCGTATGATGATGAAGTGAATGATGATGATGATGATGATGATGATGATGAT 1604
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1141 AGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1605 GAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1664
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1201 ACTGTTTGAAGAAACAAAGGTTTATGATGATGATGATGATGATGATGATGATGATG 1260
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1665 ATTCATTTGTTGGAAACAAAGGATGATGATGATGATGATGATGATGATGATGATG 1724
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1261 CATTTTCAAGATTTTACGAGAAAGATGACCGTTACTCATCTCTTCTGCTGTTGA 1320
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1725 CATCTCGAGTTTCATGAGACAAAGACTGTACAAATCCACCTCTACTATGATGTTGA 1784
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1321 GATTTGTTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1380
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1785 GGTTTTATGCTCAAACTCTGGAACCAAGCTCTACTGATGCGGAGCCATGAATGCT 1844
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 1381 GCATATCCATCTCGAGAAATGCGGTAA 1408
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 DB 1845 GCACACATTTTGAAGGCTACTGTA 1872
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13

AAA47758
 ID AAA47758 standard; cDNA; 558 BP.

AAA47758;

16-NOV-2000 (first entry)

At Hyp 2245035 (ATFC7_4) (modified) cDNA.

Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ss.

Arabidopsis thaliana.

WO20044918-A1.

03-AUG-2000.

28-JAN-2000; 2000WO-GB00248.

28-JAN-1999; 99GB-0001927.

(PLAN-) PLANT BIOSCIENCE LTD.

Dean C, Gendall A;

WPI; 2000-499333/44.

P-PSDB; AAB00065.

Isolated vernalization gene VRN2 is used to produce transgenic plants

with altered vernalization response, flowering time, leaf size and/or

shape or shade avoidance response for maximized reproductive success

Disclosure; Page 76; 105pp; English.

Isolated nucleic acid sequences obtained from the VRN2 locus of a
 plant encode polypeptides which are capable of affecting one or more
 vernalization responses such as, flowering time, leaf size and/or
 shape or the shade avoidance response of a plant into which the
 nucleic acid is introduced. Introducing such sequences into plants
 to alter these characteristics maximises the reproductive success of

CC the plant. This CDNA represents another Arabidopsis thaliana sequence
CC having homology to the VRN2 gene.
XX
SQ Sequence 558 BP; 191 A; 122 C; 112 G; 133 T; 0 other:

Query Match 13.4%; Score 230.8; DB 21; Length 558;
Best Local Similarity 69.4%; Pred. No. 3.1e-47;
Matches 361; Conservative 0; Mismatches 147; Indels 12; Gaps 3;

QY 978 AAGTATCTGCTGAGCCATGAGAGTGAAGCC---ACCTACTTCTTCGAAACGCCAA 1034
DB 31 AAGGTCCCTGCTAGCCATCAAGCTACAGCCACATCTTCCTTCAAAACGCCAG 90
QY 1035 TTCTATCTCTCAAGAGTCCAGCCATGAGCGCTTGAAGAGTAATGTGACCGGAGT 1094
DB 91 TTCTATCAATCCCGAACCGGTGACCCATGTGCTACTTGAGCAAGTATGTCTGACCGAGAT 150
QY 1095 AGCGAGATGAAGTGG-----ATGAGAGTGTGCTGAGATTTTGAAGTGGCCAGATGCTT 1148
DB 151 AGCGAATATGAGCTGACAAAATATGATGATCTGACATCTCGAAGAACCGAGATGCTT 210
QY 1149 GATGACTTTGTGAGATGATATAA---GATGAAGCAATTCATGCTCTTTGGAACTCG 1205
DB 211 AATGGTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 270
QY 1206 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1265
DB 271 TTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 330
QY 1266 TCAGATTTTACGAGAAAGAGTTCACCGTTCATGCTCTTCTGTTGGAGATTTG 1325
DB 331 TCAGATTTTACGAGAAAGAGTTCACCGTTCATGCTCTTCTGTTGGAGATTTG 1325
QY 1326 TTTTGTATTAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1385
DB 391 TTCATGATCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 450
QY 1386 ACCATCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1445
DB 451 ACCATCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 510
QY 1446 GTGATCTGCTCCAGTCTCAACACCAACCAACCAACCAACCAACCAACCAACCAACCA 1485
DB 511 ACGACTGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 550

RESULT 14
AAC54919
ID AAC54919 standard; DNA; 525 BP.
XX
AC AAC54919;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 79539.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW Protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 04-MAY-1999; 99US-0132048.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134286.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139412.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 18-JUN-1999; 99US-0139466.
PR 18-JUN-1999; 99US-0139467.
PR 21-JUN-1999; 99US-0139873.
PR 22-JUN-1999; 99US-0139877.
PR 23-JUN-1999; 99US-0139879.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140683.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.

Query	DB	Score	Length	DB	Length
Query Match		9.98	Score 170.8	DB 21	Length 525
Best Local Similarity	96.2%		Pred. No. 2.2e-32		
Matches 175	Conservative	0	Mismatches 7	Indels 0	Gaps 0
355	ACCCATGCTTCTCTCCAAAGATGCTTGAACATCAAAATTTGGAGCAAAAGCAAAAGAACT	414			
326	AGCCATGCTTCTCTCCAAAGATGCTTGAACATCAAAATTTGGAGCAAAAGCAAAAGAACT	385			
415	CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTGTAATACACATTACAGAAA	474			
386	CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTGTAATACACATTACAGAAA	445			
475	CTGAGTTTAAAGGAGATGTTCTTGTCCATTTTCTTATGCTATGTGTAGCTTCAAG	534			
446	CTGAGTTTAAAGGAGATGTTCTTGTCCATTTTCTTATGCTATGTGTAGCTTCAAG	505			
535	GG 536				
506	TG 507				

PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI: 2000-499333/44.
DR F-PSDB: AAB00064.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
XX Disclosure: Page 76; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA represents expressed sequence tag clone A1153743
CC from poplar trees (Populus tremula x Populus tremuloides) which has
CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
XX
SQ Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;

Query Match 9.5%; Score 163.8; DB 21; Length 325;
Best Local Similarity 76.4%; Pred. No. 9.8e-31;
Matches 201; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 951 GCTGTGTCCTCTACTAAGACAGAAAGTTATCTGTCAGCATCAGAGCTAGAGAC 1010
DB 62 GCCATGCTACAAATTTGCAAGACAGAAATTTATCAATGTAGCGTCTGACATGAGAAAC 121
QY 1011 CACCTACTTCTTCAGAAACGGCAATTCATCTCAGAGATCCAGCCCAATGGGCTT 1070
DB 122 CGTACACTCCCTTCACAAACGACATTTTTCACATACATAGAGCTCAGCCAAATGGCAGCT 181
QY 1071 GAGCAATTAATGTCTGACCGGATAGCGAGATGAATCGATGACATGTTCGAGATTTT 1130
DB 182 GAGCAGTTATGTGATGCGGATAGTGAGATGAAGTGAACGATGTGTCAGATTTT 241
QY 1131 GAGATGCCAGATGCTTGATGACTTTGTGATGATAAAGATGAAAAGCAATTCATG 1190
DB 242 GAAAGCCGAGAGATGCTTGATGATTTGTAGACCTGACTAAAGATGAGAAATGATG 301
QY 1191 CATCTTTGGAACCTGTTTGTAG 1213
DB 302 CACTTGTGGAACCTCATTTGTGAG 324

Search completed: June 19, 2003, 20:10:38
Job time : 430 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:56:48 ; Search time 90 Seconds

(Without alignments)
5867.749 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctcttaatttgct.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NA:*

1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*

2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*

3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*

4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*

5: /cgn2_6/ptodata/1/ina/PCtus.COMB.seq:*

6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result	No.	Score	Query	Match	Length	DB	ID	Description
C	1	69	4.0	198	5	PCT	US95-10668-3	Sequence 3, Appl1
	2	69	4.0	198	5	PCT	US95-10668-4	Sequence 4, Appl1
	3	65.8	3.8	198	5	PCT	US95-10668-1	Sequence 1, Appl1
	4	65.8	3.8	198	5	PCT	US95-10668-2	Sequence 2, Appl1
	5	62	3.6	5511	3	US	-08-928-361B-2	Sequence 2, Appl1
	6	62	3.6	7334	3	US	-08-928-361B-1	Sequence 1, Appl1
	7	60.4	3.5	5163	3	US	-08-700-651-1	Sequence 1, Appl1
	8	60.4	3.5	5163	3	US	-08-928-361B-4	Sequence 4, Appl1
	9	60.4	3.5	5318	3	US	-08-700-651-2	Sequence 2, Appl1
	10	60.4	3.5	5318	3	US	-08-928-361B-3	Sequence 3, Appl1
C	11	59.2	3.4	1086	1	US	-08-415-751-27	Sequence 27, Appl1
	12	59.2	3.4	1086	1	US	-08-415-751-28	Sequence 28, Appl1
	13	53.4	3.1	1690	1	US	-08-276-452A-24	Sequence 24, Appl1
	14	53.4	3.1	1690	2	US	-08-798-744-12	Sequence 24, Appl1
	15	51	3.0	2032	4	US	-09-241-581B-5	Sequence 5, Appl1
	16	51	3.0	2032	4	US	-08-265-428-5	Sequence 5, Appl1
	17	51	3.0	2032	5	PCT	US95-07721-5	Sequence 5, Appl1
	18	49.8	2.9	1086	1	US	-08-415-751-47	Sequence 47, Appl1
	19	48.2	2.8	688	4	US	-08-998-416-915	Sequence 915, Appl1
	20	47.6	2.8	1430	1	US	-08-276-452A-25	Sequence 25, Appl1
C	21	47.6	2.8	1430	2	US	-08-798-744-25	Sequence 25, Appl1
	22	45.8	2.7	2190	4	US	-09-625-188-19	Sequence 19, Appl1
	23	43.6	2.5	1107	2	US	-08-991-300-1	Sequence 1, Appl1
	24	43.4	2.5	3292	1	US	-07-814-964-12	Sequence 12, Appl1
	25	43.4	2.5	3292	1	US	-08-258-442-12	Sequence 12, Appl1
	26	43.4	2.5	3292	1	US	-08-328-809-7	Sequence 7, Appl1
	27	43.4	2.5	3292	5	PCT	US92-11107-12	Sequence 12, Appl1

C 28	40.8	2.4	87350	3	US-08-781-891-79	Sequence 79, Appli
C 29	40.8	2.4	87543	4	US-09-791-211-3	Sequence 3, Appli
C 30	39.8	2.3	915	4	US-09-134-001C-1741	Sequence 1741, Ap
C 31	39.6	2.3	1529	2	US-08-477-451-17	Sequence 17, Appli
C 32	39.6	2.3	1529	2	US-08-477-451-21	Sequence 21, Appli
C 33	39.6	2.3	2295	1	US-08-375-300-3	Sequence 3, Appli
C 34	39.6	2.3	2295	3	US-09-177-431-3	Sequence 3, Appli
C 35	39.6	2.3	2295	5	PCT-US95-16930-3	Sequence 3, Appli
C 36	39.6	2.3	4080	1	US-08-375-300-1	Sequence 1, Appli
C 37	39.6	2.3	4080	3	US-09-177-431-1	Sequence 1, Appli
C 38	39.6	2.3	4080	5	PCT-US95-16930-1	Sequence 1, Appli
C 39	38.8	2.3	658	4	US-08-998-416-595	Sequence 595, App
C 40	38.8	2.3	912	4	US-09-134-001C-1352	Sequence 1352, Ap
C 41	38.6	2.2	399	3	US-08-713-559-6	Sequence 6, Appli
C 42	38.4	2.2	444	4	US-09-178-509-2	Sequence 2, Appli
C 43	38.4	2.2	450	1	US-08-090-523-28	Sequence 28, Appli
C 44	38.4	2.2	450	1	US-08-398-627-28	Sequence 28, Appli
C 45	38.4	2.2	450	1	US-08-406-857-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
Sequence 3, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
TITLE OF INVENTION: Restriction Mapping
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 198
TYPE: Nucleic Acid
STRANDEDNESS: Single
TOPOLOGY: Linear
ANTI-SENSE: NO
PCT-US95-10668-3
Query Match 4.0%; Score 69; DB 5; Length 198;
Best Local Similarity 60.3%; Pred. No. 1.2e-09;
Matches 114; Conservative 0; Mismatches 75; Indels 0; Gaps 0;
1373 CAACAACGCAATGCATCCGAGAAATTCGCTAATAGTCACACACACACACACAA 1432
||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rn1

Page 2

[illegible]

RESULT 2
 PCT-US95-10668-4
 : Sequence 4, Application PC/TUS9510668
 : GENERAL INFORMATION:
 APPLICANT: James Eberwine
 TITLE OF INVENTION: A Method of Sequencing Proteins by
 TITLE OF INVENTION: Epitope Ordering and Protein
 TITLE OF INVENTION: Restriction Mapping
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ
 COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 COMPUTER: IBM 486
 OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10668
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/294,133
 FILING DATE: August 22, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: PENN-0137
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 198
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: NO
 PCT-US95-10668-4

[illegible]

Db	121	CACACACACACACACACACACACACACACACACACAA	180
QY	1553	CAAACTAAT	1561
Db	181	CAAGCTGAT	189

RESULT 3
 PCT-US95-10668-1
 Sequence 1, Application PC/TUS9510668
 GENERAL INFORMATION:
 APPLICANT: James Eberwine
 TITLE OF INVENTION: A Method of Sequencing Proteins by
 TITLE OF INVENTION: Epitope Ordering and Protein
 TITLE OF INVENTION: Restriction Mapping
 NUMBER OF SEQUENCES: 4
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Jane Massey Licata, Esq.
 STREET: 210 Lake Drive East, Suite 201
 CITY: Cherry Hill
 STATE: NJ
 COUNTRY: USA
 ZIP: 08002
 COMPUTER READABLE FORM:
 MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
 COMPUTER: IBM 486
 OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
 SOFTWARE: WORDPERFECT 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: PCT/US95/10668
 FILING DATE: Herewith
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/294,133
 FILING DATE: August 22, 1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Jane Massey Licata
 REGISTRATION NUMBER: 32,257
 REFERENCE/DOCKET NUMBER: PENN-0137
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (609) 779-2400
 TELEFAX: (609) 779-8488
 INFORMATION FOR SEQ. ID NO.: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 198
 TYPE: Nucleic Acid
 STRANDEDNESS: Single
 TOPOLOGY: Linear
 ANTI-SENSE: No
 PCT-US95-10668-1

[illegible]

```

? TITLE OF INVENTION: SPECIES INFECTIONS
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PETERS, VERNY, JONES & BIRKSA
? STREET: 385 Sherman Avenue, Suite 6
? CITY: Palo Alto
? STATE: CA
? COUNTRY: USA
? ZIP: 94306-1840
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: IBM PC compatible
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/928,361B
? FILING DATE: 12-SEP-1997
? CLASSIFICATION:
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 60/026,062
? FILING DATE: 13-SEP-1996
? ATTORNEY/AGENT INFORMATION:
? NAME: Verry, Hana
? REGISTRATION NUMBER: 30,518
? REFERENCE/DOCKET NUMBER: 480.76-1(HV)
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650-324-1677
? TELEFAX: 650-324-1678
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 511 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: double
? TOPOLOGY: linear
? MOLECULE TYPE: DNA (genomic)
US-08-928-361B-2

Query Match      3.6%; Score 62; DB 3; Length 511;
Best Local Similarity 57.1%; Pred. No. 5.2e-07;
Matches 113; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY    1358 CGACTGCGCCACCATCAGCAACATGCATTCACATCTCGAGAAATTGCCGTATAGCTCAGA 1417
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    2150 CTACTAAGGGCAACAACACCAACAACATTTTATCCAAATTAACAACAACTCAAA 2209

QY    1418 CACCACCACCAACCAACAACAACAGTGATGTCGTCAGTAGCTCAACACCAACAA 1477
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    2210 AACCAACAACAACAACAACAAGAGTTCCAGGTAAAGCCACCAATAGCCACAACACGA 2269

QY    1478 CAATAACATTTGGATCATCCCAATGACATTAACAACAAGAACAAATGTTGACAACAGA 1537
       ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db    2270 CAACATTAAGGCCAATATGTTACAAACAACAACAAGAACAAACAACAACACACAGCA 2329

QY    1538 CAATTAACGACGAGACAA 1555
       || | || | || |
Db    2330 CAGTGCCAACGACACACTA 2347

RESULT 6
US-08-928-361B-1
? Sequence 1, Application US/08928361B
? Patent No. 6071518
? GENERAL INFORMATION:
? APPLICANT: Petersen, Carolyn
? TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
? TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
? TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
? NUMBER OF SEQUENCES: 30
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PETERS, VERNY, JONES & BIRKA
? STREET: 385 Sherman Avenue, Suite 6
? CITY: Palo Alto
```


	Matches	112;	Conservative	0;	Mismatches	86;	Indels	0;	Gaps	0;
OY	1358	CGACTCAGCCACCACTCACACAGTCGAATACCATCCTGAGAAATTGGCCGTAAATAGTCA	1417							
Db	1799	CTACTTAAGGGCAACCAACCAACCAACTTTAATTCAMTCATTACAACAACAACCTCAA	1858							
OY	1418	CACCAACCAACCAACAACAACAAGTGATGATGTCGCCAGTGCTCAACACCAACAA	1477							
Db	1659	AACCAACACACACACACACAAAGTTTCCAGSTAAAGCCACCAATAGCCACAACAACAA	1918							
OY	1478	CAATAACATGTTGGATCATCCCAATGCATTAACAAACAAGAACAATGTTGAACAACAAGA	1537							
Db	1919	CACATTTAAAGCCCAATAGTTACACAACAACAACAACAAGCAACAACAACAACAACA	1978							
OY	1538	CAATTAACAGCAGAGAACAA	1555							
Db	1979	CAGTGCCACGACAACCTA	1996							

```

1      RESULT 9
2      US-08-700-651-2
3      ; Sequence 2, Application US/08700651B
4      ; Patent No. 6015882
5      ; GENERAL INFORMATION:
6      ; APPLICANT: PETERSEN, CAROLYN
7      ; APPLICANT: LEECH, JAMES
8      ; APPLICANT: NELSON, RICHARD, C.
9      ; APPLICANT: GUY, JIRI
10     ; TITLE OF INVENTION: VACCINES, ANTIBODIES, PROTEINS, GLYCOPROTEINS, DNAS AND RNAS
11     ; TITLE OF INVENTION: FOR PROPHYLAXIS AND TREATMENT OF Cryptosporidium parvum
12     ; TITLE OF INVENTION: INFECTIONS
13     ; FILE REFERENCE: 480.19-4(HV)
14     ; CURRENT APPLICATION NUMBER: US/08/700,651B
15     ; CURRENT FILING DATE: 1997-08-14
16     ; EARLIER APPLICATION NUMBER: 08/415,751
17     ; EARLIER FILING DATE: 1995-04-03
18     ; NUMBER OF SEQ ID NOS: 15
19     ; SOFTWARE: PatentIn Ver. 2.0
20     ; SEQ ID NO 2
21     ; LENGTH: 5318
22     ; TYPE: DNA
23     ; ORGANISM: Cryptosporidium parvum
24     ; US-08-700-651-2

```

	Query Match	3.5%:	Score 60.4:	DB 3:	Length 5318;
	Best Local Similarity	56.6%:	Pred. No.1.4e-06:		
	Matches 112:	Conservative	0:	Mismatches 86:	Indels 0:
				Gaps 0:	
QY	1358	CGACTCGCCACCATCATCACTTCATATACATCCTCGAGATATGCGGTATATAGCTAGA	1417		
Db	1799	CTACTAAGGGACAAACAAACCAACCACTCTTATTCACATCATTTACAAACCAACTGAAA	1858		
QY	1418	CACGACCAACCAACCAACCAACAGTGTGTGATGTCGCCAGTACTTCACACCCAA	1477		
Db	1859	AACCAACAAACCAACCAACCAACAAAGGTTCAGGTAAAGCCACCAATATGCCCAACCA	1918		
QY	1478	CAATTAACATTTGTGGATCATCCCAATGTGACATTAACCAAGACATGTTGCAACAAGA	1537		
Db	1919	CAACATTTAAAGCCCAATGTGTACAAACAACAACAAAGCAACAACAACAAACAACAA	1978		
QY	1538	CAATTAACGACGACAGACAA	1555		
Db	1979	CAGTGCCACGACAACTA	1996		

```

RESULT 10
US-08-928-361B-3
; Sequence 3, Application US/08928361B
; Patent No. 6071518
;
; GENERAL INFORMATION:
; APPLICANT: Petersen, Carolyn
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS

```

1 TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
2
3 TITLE OF INVENTION: SPECIES INFECTIONS
4
5 NUMBER OF SEQUENCES: 30
6
7 CORRESPONDENCE ADDRESS:
8 ADDRESSEE: PETERS, VERNY, JONES & BIKSA
9 STREET: 385 Sherman Avenue, Suite 6
10 CITY: Palo Alto
11
12 STATE: CA
13
14 COUNTRY: USA
15 ZIP: 94306-1840
16
17 COMPUTER READABLE FORM:
18
19 MEDIUM TYPE: Floppy disk
20
21 COMPUTER: IBM PC compatible
22 OPERATING SYSTEM: PC-DOS/MS-DOS
23 SOFTWARE: Patent In Release #1.0, Version #1.30
24
25 CURRENT APPLICATION DATA:
26 APPLICATION NUMBER: US/08/928,361B
27 FILING DATE: 12-SEP-1997
28
29 CLASSIFICATION:
30
31 PRIOR APPLICATION DATA:
32 APPLICATION NUMBER: US 60/026,062
33 FILING DATE: 13-SEP-1996
34
35 ATTORNEY/AGENT INFORMATION:
36 NAME: VERNY, Hana
37
38 REGISTRATION NUMBER: 30,518
39
40 REFERENCE/DOCKET NUMBER: 480.76-1(HV)
41
42 TELECOMMUNICATION INFORMATION:
43 TELEPHONE: 650-324-1677
44
45 TELEFAX: 650-324-1678
46
47 INFORMATION FOR SEQ ID NO: 3:
48
49 SEQUENCE CHARACTERISTICS:
50 LENGTH: 5318 base pairs
51 TYPE: nucleic acid
52 STRANDEDNESS: double
53
54 TOPOLOGY: linear
55
56 MOLECULE TYPE: DNA (genomic)
57
58 US-08-928-361B-3

[illegible]

RESULT 11
 US-08-415-751-27
 Sequence 27, Application US/08415751
 Patent No. 5643772
 GENERAL INFORMATION:
 APPLICANT: PETERSEN, CAROLYN
 APPLICANT: LEECH, JAMES
 APPLICANT: NELSON, RICHARD, C.
 APPLICANT: GUT, JIRI
 TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
 TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
 TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
 TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
 TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND

;; TITLE OF INVENTION: DIAGNOSIS AND KIT
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PHILLIPS, MOORE, LEMPLO & FINLEY
;; STREET: 385 Sherman Avenue, Suite 6
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94306-1840
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
;; COMPUTER: PC
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/415,751
;; FILING DATE: 03-Apr-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/071,880
;; FILING DATE: June 1, 1993
;; APPLICATION NUMBER: 07/891,301
;; FILING DATE: May 29, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hana Dolezalova
;; REGISTRATION NUMBER: 30,518
;; REFERENCE/DOCKET NUMBER: 480.19-2 (HND)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-1677
;; TELEFAX: (415) 324-1678
;; INFORMATION FOR SEQ ID NO: 27:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1086 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Cryptosporidium parvum
US-08-415-751-27

Query Match
Best Local Similarity 56.8%; Score 59.2; DB 1; Length 1086;
Matches 109; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1376 CAACGTGCAATACCATCTCGAGAAATTCGCGTAATAGCTCAGACACCACCACCACCA 1435
DB 341 CAACCAACACCAACCAACTCTTAATCCATCATTAACACCAACTCAAAAAACCAACA 400
QY 1436 CAACACAGTGTGATCGTCCAGTGAACCAACCAACCAATATTAATGTGATCA 1495
DB 401 CAACCAACCAACCAAGGTTCCAGGTAGCCACCAATAGCCACCAACCAACATTA 460
QY 1496 TCCCAATGACATTAACCAACCAATGTTGACACCAACCAACCAATTAACAGAGACAA 1555
DB 461 AGCCAATAGTTACCAACCAACCAACCAACCAACCAACCAACCAACCAACAGTGCCA 520
QY 1556 AGTAATTAATA 1567
DB 521 CGACAACACTACTA 532

;; TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
;; TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
;; TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
;; TITLE OF INVENTION: DIAGNOSIS AND KIT
;; NUMBER OF SEQUENCES: 50
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: PHILLIPS, MOORE, LEMPLO & FINLEY
;; STREET: 385 Sherman Avenue, Suite 6
;; CITY: Palo Alto
;; STATE: California
;; COUNTRY: United States of America
;; ZIP: 94306-1840
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
;; COMPUTER: PC
;; OPERATING SYSTEM: DOS
;; SOFTWARE: Wordperfect 5.1
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/415,751
;; FILING DATE: 03-Apr-1995
;; CLASSIFICATION: 435
;; PRIOR APPLICATION DATA:
;; APPLICATION NUMBER: 08/071,880
;; FILING DATE: June 1, 1993
;; APPLICATION NUMBER: 07/891,301
;; FILING DATE: May 29, 1992
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Hana Dolezalova
;; REGISTRATION NUMBER: 30,518
;; REFERENCE/DOCKET NUMBER: 480.19-2 (HND)
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (415) 324-1677
;; TELEFAX: (415) 324-1678
;; INFORMATION FOR SEQ ID NO: 28:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 1086 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
;; ORIGINAL SOURCE:
;; ORGANISM: Cryptosporidium parvum
US-08-415-751-28

Query Match
Best Local Similarity 56.8%; Score 59.2; DB 1; Length 1086;
Matches 109; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

QY 1376 CAACGTGCAATACCATCTCGAGAAATTCGCGTAATAGCTCAGACACCACCACCACCA 1435
DB 746 CAACCAACCAACCAACCAACTCTTAATCCATCATTAACCAACCAACTCAAAAAACCAACA 687
QY 1436 CAACACAGTGTGATCGTCCAGTGAACCAACCAACCAATATTAATGTGATCA 1495
DB 686 CAACCAACCAACCAAGGTTCCAGGTAGCCACCAATAGCCACCAACCAACCAACTTA 627
QY 1496 TCCCAATGACATTAACCAACCAATGTTGACACCAACCAACCAATTAACAGAGACAA 1555
DB 626 AGCCAATAGTTACCAACCAACCAACCAACCAACCAACCAACCAACCAACAGTGCCA 567
QY 1556 AGTAATTAATA 1567
DB 566 CGACAACACTACTA 555

RESULT 13
US-08-276-452A-24
; Sequence 24, Application US/08276452A
; Patent No. 5646029
; GENERAL INFORMATION:
; APPLICANT: Chen, Chao-Guang
; APPLICANT: Ma, Shao-Lim
; APPLICANT: Du, He

;; TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
;; TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA

APPLICANT: Gane, Allison M
 APPLICANT: Bacic, Antony
 APPLICANT: Clarke, Adrienne E
 TITLE OF INVENTION: Plant Arabinogalactan Protein (AGP) Genes
 NUMBER OF SEQUENCES: 91
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Greenlee and Winner, P.C.
 STREET: 5370 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: United States of America
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/276,452A
 FILING DATE: 18-JUL-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Caruthers, Jennie M.
 REGISTRATION NUMBER: 34,464
 REFERENCE/DOCKET NUMBER: 27-91A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303)499-8080
 TELEFAX: (303)499-8089
 TELEX: 49617824
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1690 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 60..1442
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1..38
 OTHER INFORMATION: /note="Nucleotide sequence
 OTHER INFORMATION: obtained by PCR which does not overlap with the
 OTHER INFORMATION: cDNA clone"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 60..128
 OTHER INFORMATION: /note="Predicted transmembrane
 OTHER INFORMATION: segment"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 135..179
 OTHER INFORMATION: /note="Derived amino acid sequence
 OTHER INFORMATION: corresponding to the peptide sequence by protein
 OTHER INFORMATION: microsequencing"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 135..179
 OTHER INFORMATION: /note="Amino acid 26 may also be
 OTHER INFORMATION: and 40 are identical to that in the peptide
 OTHER INFORMATION: obtained by direct microsequencing"
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 135..179
 OTHER INFORMATION: /note="Amino acid 26 may also be
 OTHER INFORMATION: and 40 are identical to that in the peptide
 OTHER INFORMATION: obtained by direct microsequencing"
 US-08-276-452A-24

Query Match 3.1%; Score 53.4; DB 1; Length 1690;
 Best Local Similarity 55.7%; Pred. No. 6.4e-05;
 Matches 102; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1373 CACCAATGCAATACCATCCTGAGATTGCCCTAATAGCTGACACCAACCAACAA 1432
 Db 920 CCAGATTACATGACAAACAAACAAATGCTTCCTCCGAGAGTTACAAACAAACAA 979
 QY 1433 CACCAACAAACAGTGTGATCGTCCAGTCACTCAACACCAACAAATTAATTTGGA 1492
 Db 980 CACCAACAAACAAACAGTGTTCCTGAGATTACAAACAAACAAATTAATTTGTTT 1039
 QY 1493 TCATCCCATGACATTAACAAACAAACAAATGTTGACAAACAGGACATTAACGAGAGA 1552
 Db 1040 CTCGAGAAATTACAAACAAACAAATTAACAAATGCTTCTAGAGAAATTACAAACAA 1099
 QY 1553 CAA 1555
 Db 1100 CAA 1102

RESULT 14
 US-08-798-744-24
 Sequence 24, Application US/08798744
 Patent No. 5830747
 GENERAL INFORMATION:
 APPLICANT: Chen, Chao-Guang
 APPLICANT: Mau, Shiao-Lim
 APPLICANT: Du, He
 APPLICANT: Gane, Allison M
 APPLICANT: Bacic, Antony
 APPLICANT: Clarke, Adrienne E
 TITLE OF INVENTION: Plant Arabinogalactan Protein (AGP) Genes
 NUMBER OF SEQUENCES: 91
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Greenlee and Winner, P.C.
 STREET: 5370 Manhattan Circle, Suite 201
 CITY: Boulder
 STATE: Colorado
 COUNTRY: United States of America
 ZIP: 80303
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: IBM PC compatible
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/798,744
 FILING DATE: 13-FEB-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/276,452
 FILING DATE: 18-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Caruthers, Jennie M.
 REGISTRATION NUMBER: 34,464
 REFERENCE/DOCKET NUMBER: 27-91A
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (303)499-8080
 TELEFAX: (303)499-8089
 TELEX: 49617824
 INFORMATION FOR SEQ ID NO: 24:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1690 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 60..1442
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: 1..38
 OTHER INFORMATION: /note="Nucleotide sequence
 OTHER INFORMATION: obtained by PCR which does not overlap with the
 OTHER INFORMATION: cDNA clone"

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rn1

Page 8

```
FEATURE:
NAME/KEY: misc.feature
LOCATION: 60-128
OTHER INFORMATION: /note= "Predicted transmembrane
OTHER INFORMATION: segment"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note= "Derived amino acid sequence
OTHER INFORMATION: corresponding to the peptide sequence by protein
OTHER INFORMATION: microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note= "Amino acids 27 to 36, 38,
OTHER INFORMATION: and 40 are identical to that in the peptide
OTHER INFORMATION: obtained by direct microsequencing"
FEATURE:
NAME/KEY: misc.feature
LOCATION: 135-179
OTHER INFORMATION: /note= "Amino acid 26 may also be
OTHER INFORMATION: Ala; 37 and 39 can also be undetermined residues"
US-08-798-744-24
```

```
Query Match 3.1%; Score 53.4; DB 2; Length 1690;
Best Local Similarity 55.7%; Pred. No. 6.4e-05;
Matches 102; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1373 CAACAACCTGCAATACCATCTCGAATTCGCGTATAGCTCAGACACACACACACAA 1432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 920 CCGAATTCATGAAACAACAACATGCGCTTCGAGAGTTACAACAACAACAACA 979
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1433 CAACAACAAGTGTGATGCTCCAGTCAACACCAACAACAACAATATGTTGGA 1492
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 980 CACAACAACAACAACGTTTCTCTGGAATTAACAACAACAACAATATGATGTTT 1039
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1493 TCATCCCATGACATAAACAACAACAATATGACAACAAGACAAATATACGACGAGA 1552
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1040 CTCGGAATTCACAACAACAACAATATGACAAATGCTTTCACGAGATTACAACAACA 1099
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1553 CAA 1555
    |||
DB 1100 CAA 1102
```

```
RESULT 15
US-09-241-581B-5
; Sequence 5, Application US/09241581B
; Patent No. 6350859
; GENERAL INFORMATION:
; APPLICANT: Massachusetts Institute of Technology
; TITLE OF INVENTION: Class BI and CI Scavenger Receptors
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/241,581B
; FILING DATE: 02-Feb-1999
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
```

```
REFERENCE/DOCKET NUMBER: MTT6620
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404) 873-8794
TELEFAX: (404) 873-8795
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 2032 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: misc.feature
LOCATION: 40-1926
OTHER INFORMATION: /Function = "Nucleotides 40 through
1926 encode the amino acid sequence
for the Drosophila Melanogaster
Scavenger Receptor Class CI."
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-241-581B-5
```

```
Query Match 3.0%; Score 51; DB 4; Length 2032;
Best Local Similarity 57.4%; Pred. No. 0.00032;
Matches 112; Conservative 0; Mismatches 80; Indels 3; Gaps 1;

QY 1358 CGACTGACCCACCATCAACAACATGCAATACCATCTCGAATTCGCGTATAGCTCAGA 1417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1346 CAACAAGGCGCAACACTCAAAAGGAACAACAACACTTAAAAACGACAAACAACCTCAA 1405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1418 CACGACGACGACCAACAACAACAAGTGTGATGCTCCAGTCACTCAAAACCAACA 1477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1406 CACGCGGACGACCAACAACAAGTGTGATGCTCCAGTCACTCAAAACCTCACA 1462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1478 CAATACATTTGTGATCATCCATGACATTAACAACAAGAAATGTTGACAACAAGA 1537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1463 CGCTACACTTCAACAACAACAAGCAACAACAATCAATTAATGTGTTACAACAAGA 1522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1538 CAATACGACGAGA 1552
    ||| ||| |||
DB 1523 AAACAACAATATGA 1537
    ||| ||| |||
```

Search completed: June 19, 2003, 22:10:18
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 21:25:52 ; Search time 285 Seconds

(without alignments)
8866.349 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagcttcctcaattgct.....aatgattcgtctataact 1722

Scoring table: IDENTITY NUC

Gapop 10.0, Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCIT_NEM_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US07_NEM_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US09_NEM_PUB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	187.2	10.9	2556	US-09-938-842A-2418	Sequence 2418, App
2	92.4	5.4	888	US-09-938-842A-453	Sequence 453, App
3	55.8	3.2	2886	US-09-801-368-131	Sequence 131, App
4	54.8	3.2	4985	US-10-094-240-10	Sequence 10, Appl
5	52	3.0	1236	US-10-077-584-3	Sequence 3, Appl
6	51.6	3.0	867	US-09-216-393-340	Sequence 340, App
7	51.6	3.0	867	US-09-216-393-342	Sequence 342, App
8	51.6	3.0	1397	US-09-216-393-345	Sequence 345, App
9	51.6	3.0	1397	US-09-216-393-345	Sequence 345, App
10	51.4	3.0	333	US-09-938-842A-1152	Sequence 3152, App
11	45.8	2.7	2901	US-09-801-368-371	Sequence 371, App
12	45.2	2.6	423	US-09-918-995-7147	Sequence 7147, App
13	44.8	2.6	2277	US-10-026-188-6	Sequence 6, Appl
14	44.4	2.6	155074	US-09-801-368-223	Sequence 223, App
15	43.4	2.5	1560	US-09-893-519A-82	Sequence 82, Appl
16	43	2.5	1257	US-09-822-830A-428	Sequence 428, App
17	43	2.5	2217	US-09-874-162A-6	Sequence 6, Appl
18	43	2.5	2328	US-09-874-162A-9	Sequence 9, Appl
19	43	2.5	2481	US-09-764-864-124	Sequence 124, App

20	43	2.5	3150	US-09-938-842A-243	Sequence 243, App
21	43	2.5	3545	US-09-764-864-144	Sequence 144, App
22	43	2.5	4409	US-09-874-162A-7	Sequence 7, Appl
23	43	2.5	4441	US-09-874-162A-4	Sequence 4, Appl
24	43	2.5	4441	US-09-969-347-169	Sequence 169, App
25	43	2.5	6030	US-10-239-676-164	Sequence 164, App
26	42.6	2.5	456	US-09-864-761-1429	Sequence 4249, App
27	42.2	2.5	58985	US-09-901-153-3	Sequence 3, Appl
28	41.8	2.4	500	US-09-991-936-531	Sequence 521, App
29	41.8	2.4	5085	US-10-198-846-9854	Sequence 9854, App
30	41.4	2.4	493	US-09-864-761-5863	Sequence 5863, App
31	41.2	2.4	14649	US-10-239-676-122	Sequence 122, App
32	41	2.4	1530	US-09-759-777-113	Sequence 113, App
33	40.8	2.4	779	US-09-772-134B-69	Sequence 69, Appl
34	40.8	2.4	2614	US-09-822-846-491	Sequence 491, Appl
35	40.8	2.4	6167	US-10-239-676-44	Sequence 42, Appl
36	40.8	2.4	6604	US-09-880-107-1748	Sequence 1748, App
37	40.8	2.4	335913	US-09-754-853A-2	Sequence 2, Appl
38	40.8	2.4	335913	US-09-754-853A-3	Sequence 3, Appl
39	40.6	2.4	1075	US-09-864-761-19241	Sequence 19241, A
40	40.6	2.4	1403	US-09-864-761-2513	Sequence 2513, App
41	40.4	2.3	2031	US-09-801-368-301	Sequence 301, App
42	40.4	2.3	1830121	US-10-329-960-1	Sequence 1, Appl
43	40.2	2.3	411	US-09-960-352-14521	Sequence 14521, A
44	40.2	2.3	17419	US-10-239-676-100	Sequence 100, App
45	40	2.3	6306	US-10-239-676-130	Sequence 130, App

ALIGNMENTS

RESULT 1
US-09-938-842A-2418
Sequence 2418, Application US/09938842A
Patent No. US2002016037B1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
TITLE OF INVENTION: SAME, AND METHODS OF USE
FILE REFERENCE: SCRIPT000-3
CURRENT APPLICATION NUMBER: US/09/938,842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2418
LENGTH: 2556
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2418

Query Match 10.9%; Score 187.2; DB 9; Length 2556;
Best Local Similarity 89.7%; Pred. 2.8e-37;
Matches 201; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

816 AATGATGAACCGGTTTGATATCCCGGCAACAGAGCTTGCGGCAATTGAG 875
|||||
2332 ACTGAGGAAACCGGTGTTAGATATCCCGGCAACAGAGCTTGCGGCAATTGAG 2391
|||||
876 ATGACCAACCAATTCACACAGCCATFACCCACTTCTTGAGAGCTGCTGCTAAATT 935
|||||
2392 ATGACTACCAATTCACACAGCCATFACCCACTTCTTGAGAGCTGCTGCTAAATT 2451
|||||
936 ATATTGACACAGCAAGCTGTGTCCCTCTACTAGACAGAAAGTTATCTGCTGAGCGA 995
|||||
2452 ATATTGACACAGCAAGCTGTGTCCCTCTACTAGACAGAAAGTTATCTGCTGAGCGA 2511
|||||

OY 996 TCAGAGGCTAGAACCCACTTCTTCAGAAAGCCATCTTA 1039
|||||
Db 2512 TCAGAGGCTAGAAAGGTTTTCATCATCATGACCCCGCTATCATTA 2555

RESULT 2

US-09-938-842A-453
Sequence 453, Application US/0993842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
APPLICANT: Zhu, Tong
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCRIPI300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227,866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264,647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 453
LENGTH: 888
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-453

Query Match 5.4%; Score 92.4; DB 9; Length 888;
Best Local Similarity 73.4%; Pred. No. 2.4e-13;
Matches 160; Conservative 0; Mismatches 1; Indels 57; Gaps 1;

OY 1062 ATGGCGCTGACGAGTAATGTCTGACCGGATAGGAGATGAAGTGCATGTT 1121
|||||
Db 1 ATGGCGCTGACGAGTAATGTCTGATCGGATAGGAGATGAAGTGCATGTT 60
OY 1122 GCAGATTTGAAGATGCCAG----- 1142
|||||
Db 61 GCAGATTTGAAGATGCCAGTATTCATGATTTCTGCGTCAATTAAGTAGCAA 120
OY 1143 -----ATCCTTGATGACTTGTGATGTAATAAAGATGAAGCAA 1184
|||||
Db 121 CAGAAATGATATGATATGATATCTGATGATCTTGTGATGTAATAAAGATGAAGCAA 180
OY 1185 TTCATGATCTTTGGAAGTCTGTTGTGAAGAAACAAG 1222
|||||
Db 181 TTCATGATCTTTGGAAGTCTTGTGAAGAAACAAG 218

RESULT 3

US-09-801-368-131
Sequence 131, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Calli, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801.368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: PatentIn version 3.0
SEQ ID NO 131
LENGTH: 2886
TYPE: DNA
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-131

Query Match 3.2%; Score 55.8; DB 10; Length 2886;
Best Local Similarity 54.0%; Pred. No. 0.0011;
Matches 114; Conservative 0; Mismatches 97; Indels 0; Gaps 0;

OY 1364 AGCCGATGATACACTGATACATACCTCTCGAATTCCTTAATAGCTACAGACAC 1423
|||||
Db 1461 AGCCGATGATACACTGATACATACCTCTCGAATTCCTTAATAGCTACAGACAC 1520
OY 1424 CACCACCAACACACACACAGTGTGATCGTCCAGTGCATCAACACCAACAAATA 1483
Db 1521 CGATTAACGATTAACATTAATAGCAATTAATTAATTAATTAATTAATTAATTA 1580
OY 1484 CATTGTGATCATCCCATGACATTAACACACAGAAATGTTGACACAAAGCAATA 1543
Db 1581 TAATTAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1640
OY 1544 CAGCAGACACAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1574
Db 1641 TATTAATAAATAAGTTGATTAACATTAATTAATTAATTAATTAATTAATTAATTA 1671

RESULT 4

US-10-094-240-10/c
Sequence 10, Application US/10094240
Publication No. US20030082637A1
GENERAL INFORMATION:
APPLICANT: ZWIBEL, LAURENCE J.
TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
CURRENT APPLICATION NUMBER: US/10/094,240
CURRENT FILING DATE: 2001-03-08
PRIOR APPLICATION NUMBER: 10/056,405
PRIOR FILING DATE: 2002-01-24
PRIOR APPLICATION NUMBER: 60/264,649
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 27
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 10
LENGTH: 4985
TYPE: DNA
ORGANISM: Anopheles gambiae
US-10-094-240-10

Query Match 3.2%; Score 54.8; DB 9; Length 4985;
Best Local Similarity 47.6%; Pred. No. 0.0028;
Matches 161; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

OY 1383 AATACCATCTCGAATTTGGCGTAATAGTCAACACACACACACACACACAC 1442
|||||
Db 3063 AATACCATCTCGAATTTGGCGTAATAGTCAACACACACACACACACACAC 3004
OY 1443 AGTGTGATCGTCCAGTCACTCAACACACACACATTAATTTGATCAATCCCAAT 1502
Db 3003 AATACCATCTCGAATTTGGCGTAATAGTCAACACACACACATTAATTTGATCAAT 2944
OY 1503 GACATTAACACACACACACATTTGACACACACACACACACACACACACACAC 1562
Db 2943 GACATTAACACACACACATTTGACACACACACACACACACACACACACAC 2884

[illegible]

RESULT 5

```

? US-10-077-584-3
? Sequence 3, Application US/10077584
? Publication No. US20030073610A1
? GENERAL INFORMATION:
? APPLICANT: LINDQUIST, SUSAN
? APPLICANT: KROBITSCH, SYLVIA
? APPLICANT: OOTEIRO, TINGO F.
? TITLE OF INVENTION: YEAST SCREENS FOR THE TREATMENT OF HUMAN DISEASE
? FILE REFERENCE: ARCD:367US
? CURRENT APPLICATION NUMBER: US/10/077,584
? CURRENT FILING DATE: 2002-02-15
? PRIOR APPLICATION NUMBER: 60/269,157
? PRIOR FILING DATE: 2001-02-15
? NUMBER OF SEQ ID NOS: 9
? SOFTWARE: PatentIn Ver. 2.1
? SEQ ID NO 3
? LENGTH: 1236
? TYPE: DNA
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(513)
? US-10-077-584-3

```

Query Match 3.08; Score 52; DB 9; Length 1236;

QY	1367	CACCATCAACAACGCAATACCATTCGCGAGAAATTTGGCTAATAGCTCAGACACACAC	1426
Db	58	CAGCAACAAGCAACGACGACCAACGCAACAGCAGCAACAACACAGCGAA	117
QY	1427	CACCAACAACAACACAGTGTGATCTCCAGTCAACACCAACAACATAACAT	1486
Db	118	CAGCAACAACAAGCGCAACAACGACAAACAGAGCAACAACAGCAGCAACAGCA	177
QY	1487	TGTGTGATCATCCCAATGACATTAACAACAAGAACATGTTGACACAAGGACAATAACAG	1546
Db	178	CAACAGCAGCAACGACAACAACAGCAACAACAGCAACAACAGCAGCAACAACAACAG	237
QY	1547	CAGAGACA	1554
Db	238	CAGCAACA	245

RESULT 6

```

US-09-216-393-340
: Sequence 340, Application US/09216393
: Patent No. US2001001447A1
:
: GENERAL INFORMATION:
: APPLICANT: Milhausen, Michael James
: TITLE OF INVENTION: TOKOPLASMA GONDIID PROTEINS, NUCLEIC ACID MOLECULES, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: TX-1-C2
: CURRENT APPLICATION NUMBER: US/09/216,393
: CURRENT FILING DATE: 1998-12-18
: EARLIER APPLICATION NUMBER: 08/994,825
: EARLIER FILING DATE: 1997-12-19

```

```

? NUMBER OF SEQ ID NOS: 364
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 340
? LENGTH: 867
? TYPE: DNA
? ORGANISM: Toxoplasma gondii
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(867)
US-09-216-393-540

```

Query Match	3.0%;	Score 51.6;	DB 10;	Length 867;
Best Local Similarity	54.8%;	Pred. No. 0.0063;		
Matches 102;	Conservative 0;	Mismatches 84;	Indels 0;	Gaps 0;

QY	1370	CATCAACAACGCAATATACATCTCCGGAATTTGGCGTAATGTGACGACACACACAC	1422
Db	485	CATCAACGACCACTACTCTCCACTACACCACTACGACTACTCTTCAACTCTGACACCCA	544
QY	1430	CAACAACAACAAGTGTGATCGTCCAGTGACTCAAAACCAACAACAATTAACATTGT	1489
Db	545	CAACAACGCAACAACCAACAACATACCAACAACAACGAACAACAACCAACAACCTA	604
QY	1490	GGATCATCCCATGTACATTAACAACAAGAACAAATGTTGACAACAAGCAATTAACGCG	1549
Db	605	CACCAACAACAACGACCAACAACCAACCAACTACACCAACAACAACGCAACAACAACCA	664
QY	1550	AGACAA	1555
Db	665	CGCCAA	670

RESULT 7

US-09-216-393-342
 US-09-216-393-342 Application US/09216393
 Patent No. US20010014447A1
 GENERAL INFORMATION:
 APPLICANT: Milhausen, Michael James
 TITLE OF INVENTION: TOXOPLASMA GONDII PROTEINS, NUCLEIC ACID MOLECULES, AND
 TITLE OF INVENTION: USES THEREOF
 FILE REFERENCE: TX-1-C2
 CURRENT APPLICATION NUMBER: US/09/216,393
 CURRENT FILING DATE: 1998-12-18
 EARLIER APPLICATION NUMBER: 08/994,825
 EARLIER FILING DATE: 1997-12-19
 NUMBER OF SEQ ID NOS: 364
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 342
 LENGTH: 867
 TYPE: DNA
 ORGANISM: Toxoplasma gondii
 US-09-216-393-342

Query Match	3.0%;	Score 51.6;	DB 10;	Length 867;
Best Local Similarity	54.8%;	Pred. No. 0.0063;		
Matches 102; Conservative	0;	Mismatches 84;	Indels 0;	Gaps 0;

QY	1370	CATTAACAACATGCAGTAATACATCTCTGGAGAAATTGGCCGTATATAGCTCAGACACCAACCAAC	1423
Db	485	CATCAACTTACCCACTACTACCACTCCACACACTATAGACTACTACTCAACTCAGAACCCAA	544
QY	1430	CACCAACAACAACAGTGGATGGATGCCAGTACTCAACACCAACAACATTAACATTTGT	1489
Db	545	CACCAACTTACACACACCAACCAACTTACACCCACAACAACGACACACACACCAACCAACACTA	604
QY	1490	GGATCATCCCAATGACATTAACAACAAGAACAAATGTTGACCAACAAGACATAAACAAG	1549
Db	605	CACCAACAACAAGACACAACCAACCAACAACTACACCAACCAACAACGACAAACCAACCA	664
QY	1550	AGACAA	1555
Db	665	CGCCAA	670


```

; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 371
; LENGTH: 2901
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-371

```

```

Query Match      2.7%; Score 45.8; DB 10; Length 2901;
Best Local Similarity 59.7%; Pred. No. 0.41;
Matches 77; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

```

```

OY 1427 CACCAACAACAACAGTGTGATCTCCAGTGCATCAACACCAACCAATTAACAT 1486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1633 CAAGCAGCAGCAAGCAAGCAAGCAGCGCAGCAACAACAACAACAAGCA 1692
OY 1487 TGTGATCATCCATGACATTAACAACAACAATGTGACACAAGCAATTAACAG 1546
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 1693 CAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAAG 1752
OY 1547 CAGACACAA 1555
    ||| |||
DB 1753 CAGCAGCAA 1761

```

```

RESULT 12
US-09-918-995-7147
; Sequence 7147, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7147
; LENGTH: 423
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (1)...(423)
; OTHER INFORMATION: n = A,T,C or G
US-09-918-995-7147

```

```

Query Match      2.6%; Score 45.2; DB 9; Length 423;
Best Local Similarity 49.7%; Pred. No. 0.18;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

```

OY 1466 AAACACAACAACAATGATGATCATCCCATGATTAACAACAACAATGAT 1525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 215 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274
OY 1526 TGACACAAGACAGATTAACAGACAAAGTAATTAATAGAAAATCTCCGCTTT 1585
    | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 275 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAANNNNNCCCTT 334
OY 1586 ATGATACCATTTATCGGATTTGTAATCTCTCTTTTAAATAATGTTT 1638
    | | | | | | | | | | | | | | | | | | | | | | | | |
DB 335 TTTTNTTTTNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 387

```

```

RESULT 13
US-09-801-368-223
; Sequence 223, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Call, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 223
; LENGTH: 2277
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-223

```

```

Query Match      2.6%; Score 44.8; DB 10; Length 2277;
Best Local Similarity 52.7%; Pred. No. 0.63;
Matches 97; Conservative 0; Mismatches 87; Indels 0; Gaps 0;

```

```

OY 1373 CAACAATGCAATACCATCTCTGAGATTCGCGTAATAGCTCAGACACCACGACAA 1432
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 805 CAATCCAGCAGACAGACCCAGCAATCATCTCAACAGAGATCCAGCAGCATCAA 864
OY 1433 CAACAACAAGTGTGATCTGCTCCAGTACTCAACACCAACAATTAATGTGGA 1492
    | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 865 CCCAGCATCAACGCCAACAGCAACGACAGCAAGCAACAACAACAACAACAAGCA 924
OY 1493 TCATCCCAATGATTAACAACAAGAACATGTTGACAACAAGCAATTAACAGAGGA 1552
    | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 925 CAAGCAGCAAGCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 984
OY 1553 CAAA 1556
    |||
DB 985 CAGA 988

```

```

RESULT 14
US-10-026-188-6/c
; Sequence 6, Application US/10026188
; Patent No. US2002016465A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Zhang, Yifeng
; TITLE OF INVENTION: The Regents of the University of California
; TITLE OF INVENTION: Assays for Taste Receptor Cell Specific
; FILE REFERENCE: 02307E-114910US
; CURRENT APPLICATION NUMBER: US/10/026,188
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US 60/259,379
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 8

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 155074
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human genomic region containing 11p15.5 PAC clone pdj915f1
; OTHER INFORMATION: (Human Chromosome 11p15.5 PAC clone pdj915f1)
; US-10-026-188-6

```

Query Match
Best Local Similarity 59.7%; Pred. No. 14;
Matches 74; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

```

QY 1366 CCACCATCAACAGTCAATTCCTGAGAAATGCGGTAAATGCTCAGACACCA 1425
Db 45767 CCACCATCAACAGTCAATTCCTGAGAAATGCGGTAAATGCTCAGACACCA 1425
QY 1426 CCACCAACACACAGTGTGATGCTCCAGTACTCAACACCAACATTAACA 45708
Db 45707 CCACCAACACACAGTGTGATGCTCCAGTACTCAACACCAACATTAACA 1485
QY 1486 TTGT 1489
Db 45647 CTGT 45644

```

```

RESULT 15
US-09-893-519A-82
; Sequence 82, Application US/09893519A
; Publication No. US20030027243A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS PHARMACEUTICALS, INC.
; APPLICANT: THOMPSON, Craig
; APPLICANT: MOORE, Jeffrey
; APPLICANT: BURMAN, Ed T.
; APPLICANT: BRADLEY, John
; APPLICANT: DESILVA, Thamara
; APPLICANT: HARRIS, Sandra
; APPLICANT: KOMARNITSKY, Svetlana
; APPLICANT: MONDILLO, Marc
; APPLICANT: MOORE, Daniel
; APPLICANT: SACOT, Melissa
; APPLICANT: SANDERSON, Karen
; APPLICANT: HNO, Tatig
; APPLICANT: ZHU, Shuhao
; APPLICANT: LONG, Fan
; APPLICANT: DAYTOO, Eugene
; TITLE OF INVENTION: ANTIFUNGAL COMPOUNDS AND METHODS OF USE
; FILE REFERENCE: 0342/16348-US2
; CURRENT APPLICATION NUMBER: US/09/893,519A
; PRIOR FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,164
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: US 60/224,457
; PRIOR FILING DATE: 2000-08-10
; NUMBER OF SEQ ID NOS: 146
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 82
; LENGTH: 1560
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: Genbank/CAA96830.1
; DATABASE ENTRY DATE: 1997-08-11
; RELEVANT RESIDUES: (1)..(1560)
US-09-893-519A-82

```

Query Match
Best Local Similarity 2.5%; Score 43.4; DB 9; Length 1560;
Matches 110; Conservative 0; Mismatches 111; Indels 0; Gaps 0;

```

QY 1330 TGATTAACATGAGAACATGAGCTTGTGACTGACCCACCATCAACAACTGCAATACCA 1389
Db 285 TGCGCAAGCTTGAGACATGAGATATGCGCAACAGCAACAGCAACAAACCAACA 344
QY 1390 TCCGAGATTTGCGGTATAGCTCAGACACCAACCAACCAACCAAGTGTGG 1449
Db 345 GCAACCAAGACATGGGCAACAGCAACCTCAACAGCACTCAACAGCACTCAACAGCA 404
QY 1450 ATGTCAGAGTACTCAACACCAACCAACCAACCAATGATGATCAATCAATGACATTA 1509
Db 405 ACCTCAACAGCAACCTCAACAGCAACCTCAACAGCAACCTCAACAGCAACCTCAACAGCA 464
QY 1510 ACAACAGCAACATGTTGACAAACAGCAACATTAACAGCA 1550
Db 465 ACCTCAACATTCACCACTTCAGCCACCAACATGAGGACCCAGA 505

```

Search completed: June 19, 2003, 23:44:22
Job time: 291 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:03:46 ; Search time 4023 Seconds
(Without alignments)
10761.960 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722
Sequence: 1 caagctcttcatttgct.....aatgattctgtataact 1722

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 24791104 segs, 12571243825 residues

Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

```

Pending_Patents_NA_Main:*
1: /cgn2_6/ptodata/1/pna/PCUS_COMB.seq:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*

```

Result No.	Score	Query Match	Length	DB ID	Description
1	1722	100.0	1722	33 US-09-890-220-1	Sequence 1, Appl1
2	1722	100.0	1722	33 US-09-890-475-57	Sequence 57, Appl1
3	1720.4	99.9	1722	33 US-09-890-220-47	Sequence 47, Appl1
4	1569.6	91.1	1715	33 US-09-890-220-4	Sequence 4, Appl1
5	1569.6	91.1	1715	33 US-09-890-475-58	Sequence 58, Appl1
6	1537.6	89.3	1737	33 US-09-890-220-7	Sequence 7, Appl1
7	979.8	56.9	1497	19 US-09-513-996A-69038	Sequence 69038, A
8	412.2	23.9	6338	33 US-09-890-220-3	Sequence 3, Appl1
9	366.8	21.3	1722	42 US-10-219-999-5373	Sequence 5373, A
10	326.4	19.0	606	22 US-09-565-309A-32201	Sequence 32201, A
11	319.4	18.5	5895	33 US-09-890-220-6	Sequence 6, Appl1
12	319.4	18.5	201471	20 US-09-534-859-587	Sequence 587, App
13	319.4	18.5	201471	31 US-09-803-736-587	Sequence 587, App
14	314.2	18.2	319	19 US-09-513-996A-69042	Sequence 69042, A
15	314	18.2	2346	74 US-60-303-460-9	Sequence 52491, A
16	306.8	17.8	638	22 US-09-565-309A-52491	Sequence 32200, A
17	306.8	17.8	640	22 US-09-565-309A-32200	Sequence 61982, A
18	306.8	17.8	640	22 US-09-565-309A-61982	Sequence 945, App
19	306.4	17.8	640	19 US-09-595-329A-945	Sequence 41806, A
20	306.4	17.8	640	19 US-09-513-996A-41806	Sequence 2, Appl1
21	291.4	16.9	2280	17 US-09-339-947A-2	

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

```

22 284.2 16.5 2249 74 US-60-303-460-1 Sequence 1, Appl1
23 264.8 15.4 2145 74 US-60-303-460-13 Sequence 13, Appl
24 257.4 14.9 1841 42 US-10-219-999-16339 Sequence 16339, A
25 257.4 14.9 1841 75 US-60-312-544-4687 Sequence 4687, Ap
26 257.4 14.9 1856 33 US-09-865-439A-20106 Sequence 20106, A
27 257.4 14.8 2459 74 US-60-303-460-5 Sequence 5, Appl1
28 255 14.8 2276 74 US-60-303-460-7 Sequence 7, Appl1
29 252.4 14.7 2268 17 US-09-335-947A-9 Sequence 9, Appl1
30 247.6 14.4 353 22 US-09-565-309A-32199 Sequence 32199, A
31 243.6 14.1 2256 74 US-60-303-460-11 Sequence 11, Appl
32 242.8 14.1 677 25 US-09-654-617-26623 Sequence 26623, A
33 242.8 14.1 677 27 US-09-684-016-26623 Sequence 26623, A
34 236.2 13.7 2640 17 US-09-398-237-6 Sequence 6, Appl1
35 234 13.6 303 25 US-09-654-617-163164 Sequence 163164,
36 230.8 13.4 303 27 US-09-890-220-13 Sequence 13, Appl
37 230.8 13.4 558 33 US-09-890-220-13 Sequence 17, Appl
38 221.8 12.9 704 74 US-60-303-460-3 Sequence 3, Appl
39 214 12.4 1176 74 US-60-303-460-3 Sequence 13, Appl
40 210.8 12.2 600 43 US-10-260-238-2870 Sequence 17, Appl
41 210.8 12.2 600 61 US-60-170-912-179 Sequence 2870, Ap
42 210.8 12.2 600 76 US-60-325-448-2870 Sequence 179, Ap
43 209.8 12.2 990 25 US-09-654-617-386601 Sequence 386601,
44 209.8 12.2 990 27 US-09-684-016-386601 Sequence 386601,
45 209 12.1 531 33 US-09-874-708A-29223 Sequence 29223, A

```

ALIGNMENTS

RESULT 1
US-09-890-220-1
Sequence 1, Application US/09890220

GENERAL INFORMATION:
APPLICANT: Dean, Caroline
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: Newburn
CURRENT APPLICATION NUMBER: US/09/890, 220
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patentm Ver. 2.1
SEQ ID NO 1722
LENGTH: 1722
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-1

Query Match 100.0%; Score 1722; DB 33; Length 1722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 CAAGCTCTCAATTTGCTGCTCTCTTACACAGCAATCGGTGTTTGCACCTT 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CAAGCTCTCAATTTGCTGCTCTCTTACACAGCAATCGGTGTTTGCACCTT 60
QY 61 CAGGCTCAATCCAAACATCTATATAGCATATGCAAGAGGGGGTTCATTTG 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 CAGGCTCAATCCAAACATCTATATAGCATATGCAAGAGGGGGTTCATTTG 120
QY 121 GCATTGAGTTATCGCTATGACGTAGGGAATTCATTTAGGGAGGCTCAGACTTG 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GCATTGAGTTATCGCTATGACGTAGGGAATTCATTTAGGGAGGCTCAGACTTG 180
QY 181 CACTTAATCATATGAGCTTGAAGTGTGATTAATTGAACAAGATGTAGGC 240
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 CACTTAATCATATGAGCTTGAAGTGTGATTAATTGAACAAGATGTAGGC 240
QY 241 AGAATGTGCGGGAATCCTCACCGGAGGAAGATTTCACTGATGAGATCTCTTA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

Db 241 AGAATGTGCGGGAATCCTCACCGGAGGAAGATTTCACTGATGAGATCTCTTA 300
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 301 TATATTTGAACCTGTTCCACATATATACATCTTTACCTTGGCTCTTGAACCAT 360
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 301 TATATTTGAACCTGTTCCACATATATACATCTTTACCTTGGCTCTTGAACCAT 360
QY 361 CGTTTCTCCAAAGTCTTGAACATCAAAATTTGAGCAACCGCAAAAGTCAAGT 420
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 361 CGTTTCTCCAAAGTCTTGAACATCAAAATTTGAGCAACCGCAAAAGTCAAGT 420
QY 421 CTACTGGATGATGATTTTCACTATATAGATTTGAATACATTCACAGAAACTGA 480
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 421 CTACTGGATGATGATTTTCACTATATAGATTTGAATACATTCACAGAAACTGA 480
QY 481 TTAGGAGAGATGTTCTTGTCCATTTTGTCTATGCTATGAGTTCAGAGGCTGC 540
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 481 TTAGGAGAGATGTTCTTGTCCATTTTGTCTATGCTATGAGTTCAGAGGCTGC 540
QY 541 AATTTCATTTGAATCATCTCATGATTTATTTGAATTTGATTCAGCTTTTGAAGAT 600
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 541 AATTTCATTTGAATCATCTCATGATTTATTTGAATTTGATTCAGCTTTTGAAGAT 600
QY 601 ACCAGACATTAATGTTTCTGTAAACTTAATCTTCAATTTTGAGAGAAAGCAAGT 660
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 601 ACCAGACATTAATGTTTCTGTAAACTTAATCTTCAATTTTGAGAGAAAGCAAGT 660
QY 661 ATGAGATAAATTTGAGCCCTCTCTCTGCTCGAACCCTGTAACGGAGCAAAAG 720
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 661 ATGAGATAAATTTGAGCCCTCTCTCTGCTCGAACCCTGTAACGGAGCAAAAG 720
QY 721 GTGGCAGAAATTAACACAGAGAGACTTAAGTATGCTTTTACGTTGATTCACCAT 780
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 721 GTGGCAGAAATTAACACAGAGAGACTTAAGTATGCTTTTACGTTGATTCACCAT 780
QY 781 TAACATATGAGCAGAAATGAAATGAAATACCTTCAATGATGAGAAACCGGTTAGAT 840
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 781 TAACATATGAGCAGAAATGAAATGAAATACCTTCAATGATGAGAAACCGGTTAGAT 840
QY 841 ATCCGAGCAGAAAGCTGCTGAGCAATTTGAGATGAGCAACATTCACACAGCA 900
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 841 ATCCGAGCAGAAAGCTGCTGAGCAATTTGAGATGAGCAACATTCACACAGCA 900
QY 901 TAGCCCACTCTTCTCTGAGAGCTGCTGCTAAAGTATATGCAAGCAAGCTGTGCTC 960
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 901 TAGCCCACTCTTCTCTGAGAGCTGCTGCTAAAGTATATGCAAGCAAGCTGTGCTC 960
QY 961 CTGCTACTAAGACAGAAAGTATCTGCTGAGCGATCAGAGCTAGAGCCACTACTTC 1020
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 961 CTGCTACTAAGACAGAAAGTATCTGCTGAGCGATCAGAGCTAGAGCCACTACTTC 1020
QY 1021 TTCAAGAAACGCAATTTCTATCATCTCACAGATCCACCAATGGCGCTTGACAATGA 1080
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1021 TTCAAGAAACGCAATTTCTATCATCTCACAGATCCACCAATGGCGCTTGACAATGA 1080
QY 1081 TGTCTGACCGGAGATGAGAGATGAAGTGAAGATGAGATGCGATTTGAAGATCGCC 1140
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1081 TGTCTGACCGGAGATGAGAGATGAAGTGAAGATGAGATGCGATTTGAAGATCGCC 1140
QY 1141 AGATGCTGATGATCTTTGAGATGATGAATGAATGAATGAATGAATGAATGAATGA 1200
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1141 AGATGCTGATGATCTTTGAGATGATGAATGAATGAATGAATGAATGAATGAATGA 1200
QY 1201 ACTGCTTTGAAGAAACAAAGGTTATAGCAGATGCTATCTCTTGGGCAATGTAAG 1260
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1201 ACTGCTTTGAAGAAACAAAGGTTATAGCAGATGCTATCTCTTGGGCAATGTAAG 1260
QY 1261 CATTTTCAAGATTTTACAGAGAAAGTTCACCGTTACTCATCTCTTGTGTTGA 1320
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1261 CATTTTCAAGATTTTACAGAGAAAGTTCACCGTTACTCATCTCTTGTGTTGA 1320
QY 1321 GATTTGTTTGAATTAACATGAGAACATGAGCTGAGATCGACCATCAATCAACT 1380
| | | | | | | | | | | | | | | | | | | | | | | | | | | |

```



```

Db 1321 GATTGTTTGTGATTAACATGACACGACTGTGACTGACGACTGACGACCATCAACAACT 1380
QY 1381 GCAATACCATCTCGAGAAATGGCCCTAATAGTGCAGACACCAACCAACAAACA 1440
Db 1381 GCAATACCATCTCGAGAAATGGCCCTAATAGTGCAGACACCAACCAACAAACA 1440
QY 1441 ACAGTGTGATCGTCCAGTGCATCAACACCAACCAATATGATGATCCCA 1500
Db 1441 ACAGTGTGATCGTCCAGTGCATCAACACCAACCAATATGATGATCCCA 1500
QY 1501 ATGACATTAACACAGAAACAAATGTGACAAACAGCAATTAACAGACAGAAAGTAA 1560
Db 1501 ATGACATTAACACAGAAACAAATGTGACAAACAGCAATTAACAGACAGAAAGTAA 1560
QY 1561 TTAATATAGGAAATATCCGGCTTTATATACCATTTATCGGATTTACTTATCTTC 1620
Db 1561 TTAATATAGGAAATATCCGGCTTTATATACCATTTATCGGATTTACTTATCTTC 1620
QY 1621 TTTCTTAAATAATGTTAGGACAAACAAATTTTATATGTTAGTGTATCAACTGAT 1680
Db 1621 TTTCTTAAATAATGTTAGGACAAACAAATTTTATATGTTAGTGTATCAACTGAT 1680
QY 1681 TACATTTTATGTTAAATAATGATGATTCGTTTAACT 1722
Db 1681 TACATTTTATGTTAAATAATGATGATTCGTTTAACT 1722

```

RESULT 2

```

US-09-890-475-57
; Sequence 57, Application US/09890475
; GENERAL INFORMATION:
; APPLICANT: Johanson, Urban
; APPLICANT: West, Joanne
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,475
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/GB00/00197
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: GB 9902660.1
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patentl Ver. 2.1
; SEQ ID NO 57
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Landsberg
; OTHER INFORMATION: erecta VERN2 CDNA
US-09-890-475-57

```

```

Query Match 100.0%; Score 1722; DB 33; Length 1722;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1722; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 CAAGTCTTCAATTTGCTGCTCTCTCTACACAGCAATCGGTGTTTGGCGCTTT 60
Db 1 CAAGTCTTCAATTTGCTGCTCTCTCTACACAGCAATCGGTGTTTGGCGCTTT 60
QY 61 CAGGCTCAATCAAGACATCTATATAGCATTTTGAAGAAGAGCGCTTCTATTTGT 120
Db 61 CAGGCTCAATCAAGACATCTATATAGCATTTTGAAGAAGAGCGCTTCTATTTGT 120
QY 121 GCATTGAGTTATCGCTATGACGTAGAGGAATTTCAATTTAGGGAGCGCTCAGAGTTG 180
Db 121 GCATTGAGTTATCGCTATGACGTAGAGGAATTTCAATTTAGGGAGCGCTCAGAGTTG 180
QY 181 CACTTAAGTCAATATCGGCTCTTGACGTTGTTGATGTAATTTGAACAAGATGTAGGC 240
Db 181 CACTTAAGTCAATATCGGCTCTTGACGTTGTTGATGTAATTTGAACAAGATGTAGGC 240

```

```

QY 241 AGAATTTGCGCGGAAATCTCCACGGAGGAAGTATTTCACTGATGAAATCTCTTGA 300
Db 241 AGAATTTGCGCGGAAATCTCCACGGAGGAAGTATTTCACTGATGAAATCTCTTGA 300
QY 301 TATATTTAAACCTTTGACATATATTAACATCTTTACCTTGCCTCTAGCAACCAT 360
Db 301 TATATTTAAACCTTTGACATATATTAACATCTTTACCTTGCCTCTAGCAACCAT 360
QY 361 CGTTTCTTCCAAGATGCTTGAATCTCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAGT 420
Db 361 CGTTTCTTCCAAGATGCTTGAATCTCAAAATTTGGAGCAAGCGCAAAAGAAAGTCAAGT 420
QY 421 CTACTGGATGATTTTCACTATTAAGATTTGAATTAACATTTACAGAAATCTGAAG 480
Db 421 CTACTGGATGATTTTCACTATTAAGATTTGAATTAACATTTACAGAAATCTGAAG 480
QY 481 TTAGGAGAGATTTGTTCTGTCATTTTGCCTATGCTATGCTGATCAAGGGCGTCG 540
Db 481 TTAGGAGAGATTTGTTCTGTCATTTTGCCTATGCTGATCAAGGGCGTCG 540
QY 541 AATTTCATTTGATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTTCCAGAAAT 600
Db 541 AATTTCATTTGATTCATCTCATGATTTATTTGAATTTGAGTCAAGCTTTTCCAGAAAT 600
QY 601 ACCAGACATTAATGTTCTGTAAACTTAATCTCATATTTAGAGAAAGAGAGTG 660
Db 601 ACCAGACATTAATGTTCTGTAAACTTAATCTCATATTTAGAGAAAGAGAGTG 660
QY 661 ATGACATTAATTTGAGCCCTTCTCTCTGCTGCAAACTGTGAAGCGGAGACAAAG 720
Db 661 ATGACATTAATTTGAGCCCTTCTCTCTGCTGCAAACTGTGAAGCGGAGACAAAG 720
QY 721 GTGGAGAAATTAACACCGAGACCTTAAGATGCTTTTACCCTGATTCACCACT 780
Db 721 GTGGAGAAATTAACACCGAGACCTTAAGATGCTTTTACCCTGATTCACCACT 780
QY 781 TAACATAATGGCAGAAATGGAATCACCCTACTTAATGATGAGAAACCGTGTAGAT 840
Db 781 TAACATAATGGCAGAAATGGAATCACCCTACTTAATGATGAGAAACCGTGTAGAT 840
QY 841 ATCCGAGGCAACAGAGCTTGCTGACAAATTTGAGATGACAGCAACATTCACAGCA 900
Db 841 ATCCGAGGCAACAGAGCTTGCTGACAAATTTGAGATGACAGCAACATTCACAGCA 900
QY 901 TAGCCACTCTCTCTGACGCTGCTGCTAAAGTATTAATGACAGCAAGCTGTGCTC 960
Db 901 TAGCCACTCTCTCTGACGCTGCTGCTAAAGTATTAATGACAGCAAGCTGTGCTC 960
QY 961 CTGCTACTAAGACAGAAATGATATCTGCTGACGATCAGAGGCTGAGAGCCACTACTTC 1020
Db 961 CTGCTACTAAGACAGAAATGATATCTGCTGACGATCAGAGGCTGAGAGCCACTACTTC 1020
QY 1021 TTCAGAAACGCCAATTTCTATCTATCTCACAGAGTCCACCAATGCGCTTGAAGATTA 1080
Db 1021 TTCAGAAACGCCAATTTCTATCTATCTCACAGAGTCCACCAATGCGCTTGAAGATTA 1080
QY 1081 TGTCTGACCGGATGACAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1140
Db 1081 TGTCTGACCGGATGACAGATGGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1140
QY 1141 AGATGCTTGAATGATCTTGTGATGTAATTAAGATGAAGAAAGCAATTCATGATCTTGA 1200
Db 1141 AGATGCTTGAATGATCTTGTGATGTAATTAAGATGAAGAAAGCAATTCATGATCTTGA 1200
QY 1201 ACTGCTTTGAAGAAACAAAGGTTATAGCAGATGCTCATATCTTGGGAGATGTAAG 1260
Db 1201 ACTGCTTTGAAGAAACAAAGGTTATAGCAGATGCTCATATCTTGGGAGATGTAAG 1260
QY 1261 CATTTTCAAGATTTTACGAGAAAGATGACACCGTTACTCATCATCTTCTGTGTGGA 1320
Db 1261 CATTTTCAAGATTTTACGAGAAAGATGACACCGTTACTCATCATCTTCTGTGTGGA 1320
QY 1321 GATTGTTTGTATTAACATATGAAACATGACACTGTGACTGACGACCAATCAACAACT 1380

```

```

Db      1321 GATTGTTTGGATTTAAACATGAAACGATGAGCTGAGCTCAGCCACCACCAACACT 1380
QY      1381 GCAATACCATCTCGGAAATTCGGTAATAGCTCAGACACCACCACCACCAACACACA 1440
Db      1381 GCAATACCATCTCGGAAATTCGGTAATAGCTCAGACACCACCACCACCAACACACA 1440
QY      1441 ACAAGTGTGATCGTCCAGTGAATCAACACCAACCAACCAATTAACATTTGATGATCCCA 1500
Db      1441 ACAAGTGTGATCGTCCAGTGAATCAACACCAACCAACCAATTAACATTTGATGATCCCA 1500
QY      1501 ATGACATTAACCAACAGAAATGTTGACACAGACATTAACAGACAGACAAAGTAA 1560
Db      1501 ATGACATTAACCAACAGAAATGTTGACACAGACATTAACAGACAGACAAAGTAA 1560
QY      1561 TTAATATGAAATTCGCGCTTTATGATAGCATTTATCGATTTGATGATTTCTTC 1620
Db      1561 TTAATATGAAATTCGCGCTTTATGATAGCATTTATCGATTTGATGATTTCTTC 1620
QY      1621 TTCTTAAAAAATGTTTGAAGAGCAACAAATTTTATATGATGATGATGATGATGAT 1680
Db      1621 TTCTTAAAAAATGTTTGAAGAGCAACAAATTTTATATGATGATGATGATGATGAT 1680
QY      1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1722
Db      1681 TACATTTTATGTTAAAAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1722

```

```

RESULT 3
US-09-890-220-47
; Sequence 47, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; APPLICANT: Gendall, Anthony
; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
; FILE REFERENCE: Newbury
; CURRENT APPLICATION NUMBER: US/09/890,220
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 47
; LENGTH: 1722
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-890-220-47

```

```

Query Match          99.9%; Score 1720.4; DB 33; Length 1722;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1721; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1 CAAGCTTCTCAATTTGCTGCTCTCTCTTACACAGCCCAATGCGTTTGGCAGCTT 60
Db      1 CAAGCTTCTCAATTTGCTGCTCTCTCTTACACAGCCCAATGCGTTTGGCAGCTT 60
QY      61 CAGGCTCAATCCAAAGACATTTCTATATAGCATTTTCACAGAGGCGGTCTTAATTTG 120
Db      61 CAGGCTCAATCCAAAGACATTTCTATATAGCATTTTCACAGAGGCGGTCTTAATTTG 120
QY      121 GCATTAAGTTTATCGCTATGACGTAGGAAATTTCTAATTTAGGGGAGGCCCTAGAGTTG 180
Db      121 GCATTAAGTTTATCGCTATGACGTAGGAAATTTCTAATTTAGGGGAGGCCCTAGAGTTG 180
QY      181 CACTAACTTCATTAATGCGCTGATGAGCTGTTGATGATTAATTAACAAAGATGTAGGC 240
Db      181 CACTAACTTCATTAATGCGCTGATGAGCTGTTGATGATTAATTAACAAAGATGTAGGC 240
QY      241 AGAATGTGCGCGGAAATCTCACGGAGAGGATTTCAACTGATGATGATGATGATGAT 300
Db      241 AGAATGTGCGCGGAAATCTCACGGAGAGGATTTCAACTGATGATGATGATGATGAT 300

```

```

Db      241 AGAATGTGCGCGGAAATCTCACGGAGAGGATTTCAACTGATGATGATGATGATGAT 300
QY      301 TATATTTAAACCTGTTGCACTATATTAACATTTTACCTTGGCTCTTAGGCAACCAT 360
Db      301 TATATTTAAACCTGTTGCACTATATTAACATTTTACCTTGGCTCTTAGGCAACCAT 360
QY      361 CGTTTCTTCCAAAGATGCTTGAATACAAATTTGAGCAAAACGCAAAAGAAATCAAGAT 420
Db      361 CGTTTCTTCCAAAGATGCTTGAATACAAATTTGAGCAAAACGCAAAAGAAATCAAGAT 420
QY      421 CTACGGGATGATGATTTTCACTATTAAGATTTGATTAACATTTGATTAACAAAGATG 480
Db      421 CTACGGGATGATGATTTTCACTATTAAGATTTGATTAACATTTGATTAACAAAGATG 480
QY      481 TTAGGAGAGATGCTTCTGTCGATTTTGTCTATGCTATGCTATGCTATGCTATGCTATG 540
Db      481 TTAGGAGAGATGCTTCTGTCGATTTTGTCTATGCTATGCTATGCTATGCTATGCTATG 540
QY      541 AATTTCAATTTGATTCATCTCATGATTTTGAATTTGATTTGATTTGATTTGATTTGAT 600
Db      541 AATTTCAATTTGATTCATCTCATGATTTTGAATTTGATTTGATTTGATTTGATTTGAT 600
QY      601 ACCAGACAGTTAATGTTTCTGTAATAAATTCCTTCATATTTGAGAAAGAAAGATG 660
Db      601 ACCAGACAGTTAATGTTTCTGTAATAAATTCCTTCATATTTGAGAAAGAAAGATG 660
QY      661 ATGACGATTAATTTGAGCCCTTCTCTCTGTCGAAACCTGTAAGCGAGACAAAGATG 720
Db      661 ATGACGATTAATTTGAGCCCTTCTCTCTGTCGAAACCTGTAAGCGAGACAAAGATG 720
QY      721 GTGCGAGAAATTAACACAGAGACCTTAAGATGCTTTTATCGTTGATTCACCAAGT 780
Db      721 GTGCGAGAAATTAACACAGAGACCTTAAGATGCTTTTATCGTTGATTCACCAAGT 780
QY      781 TAACATTAAGCAGAAATTAAGATTAACATTAAGATTAAGATTAAGATTAAGATTAAGAT 840
Db      781 TAACATTAAGCAGAAATTAAGATTAACATTAAGATTAAGATTAAGATTAAGATTAAGAT 840
QY      841 ATCCGAGAGCAACAGACTTGTGCAATTTGATGATGATGATGATGATGATGATGATGAT 900
Db      841 ATCCGAGAGCAACAGACTTGTGCAATTTGATGATGATGATGATGATGATGATGATGAT 900
QY      901 TAGCCACTCTTCTGAGCGCTGCTGCTAAGATTAATTTGACAAAGGAAAGCTGTGCTC 960
Db      901 TAGCCACTCTTCTGAGCGCTGCTGCTAAGATTAATTTGACAAAGGAAAGCTGTGCTC 960
QY      961 CTGCTACTAAGACAAAGATTTATCTGCTGAGCGCATGAGAGCTAGAAAGCACTACTTC 1020
Db      961 CTGCTACTAAGACAAAGATTTATCTGCTGAGCGCATGAGAGCTAGAAAGCACTACTTC 1020
QY      1021 TTCAAGAAAGCCCAATTTCTATTCATCTCTCACAAGTCCAGCAATGCGCTTGGACAAAGTA 1080
Db      1021 TTCAAGAAAGCCCAATTTCTATTCATCTCTCACAAGTCCAGCAATGCGCTTGGACAAAGTA 1080
QY      1081 TGTCTGACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
Db      1081 TGTCTGACGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1140
QY      1141 AGATGCTTGAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
Db      1141 AGATGCTTGAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY      1201 ACTGCTTGTAAAGAAACAAAGGATTAAGCAAGATGATGATGATGATGATGATGATGATG 1260
Db      1201 ACTGCTTGTAAAGAAACAAAGGATTAAGCAAGATGATGATGATGATGATGATGATGATG 1260
QY      1261 CATTTTCAAGATTTTACGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1320
Db      1261 CATTTTCAAGATTTTACGAAAGATGATGATGATGATGATGATGATGATGATGATGATG 1320
QY      1321 GATGCTTTTGAATTAAGATTAAGCAAGATGATGATGATGATGATGATGATGATGATGAT 1380
Db      1321 GATGCTTTTGAATTAAGATTAAGCAAGATGATGATGATGATGATGATGATGATGATGAT 1380

```

QY 1381 GCAATACATCCCGGAAATTCGCGTAATAGCTCAGACACACACACACCAACCAACAA 1440
 DB 1381 GCAATACATCCCGGAAATTCGCGTAATAGCTCAGACACACACACACCAACCAACAA 1440
 QY 1441 ACAGTGTGATGTCGCCAGTGTGACTCAACACCAACCAACATTTGGATTCATCCCA 1500
 DB 1441 ACAGTGTGATGTCGCCAGTGTGACTCAACACCAACCAACATTTGGATTCATCCCA 1500
 QY 1501 ATGACATTAACACACAGAAACATTTGTACACACAGACAAATTAACAGACAGCAAAAGTAA 1560
 DB 1501 ATGACATTAACACACAGAAACATTTGTACACACAGACAAATTAACAGACAGCAAAAGTAA 1560
 QY 1561 TTAATATAGGAATATCCGGCTTTATATAGTACGATTTATGCGATTTGATTAATTTCTTC 1620
 DB 1561 TTAATATAGGAATATCCGGCTTTATATAGTACGATTTATGCGATTTGATTAATTTCTTC 1620
 QY 1621 TTTCTTAAATAATTTGTAGAGCAACAAATTTTATATGTAGTATTAACACTGAT 1680
 DB 1621 TTTCTTAAATAATTTGTAGAGCAACAAATTTTATATGTAGTATTAACACTGAT 1680
 QY 1681 TACATTTTATGTTAAAAAATGATTCGCTTATACT 1722
 DB 1681 TACATTTTATGTTAAAAAATGATTCGCTTATACT 1722

RESULT 4

US-09-890-220-4
 ; Sequence 4, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 1715
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-4

Query Match 91.1%; Score 1569.6; DB 33; Length 1715;
 Best Local Similarity 96.5%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

QY 1 CAAAGCTTCTCAATTTGCTTG--CTCTCTCTTAACACAGCAATCGGTGTTTCGACGT 58
 DB 1 CAAAGCTTCTCAATTTGCTTGCTCTCTCTTAACAGCGCAATCGGTGTTTCGACGT 60
 QY 59 TTCAGGCTTCATCAAGACATTTCTATATTAACATATTCAGAAAGCGGTTCTAATG 118
 DB 61 TTCAGGCTTCATCAAGACATTTCTATATTAACATATTCAGAAAGCGGTTCTAATG 120
 QY 119 TTCAGTGTGATTTATGCTATGAGCGTAGGGAATTTCTAATTTAGGGGAGCGCTCAGAGTT 178
 DB 121 TTCAGTGTGATTTATGAGCAATATATGAGGAATTTCTAATTTAGGGGAGCGCTCAGAGTT 180
 QY 179 TGCATTAATCTATATGCGCTCTTGACGTTGTGAGTGTATTTGAACAAGATGTGAG 238
 DB 181 TGCATTAATCTATATGCGCTCTTGACGTTGTGAGTGTATTTGAACAAGATGTGAG 240
 QY 239 GCAAGATTTGCGCGGAATCTCTCAACGAGAGAGTGTATTTCACTGATGAATCTCTT 298
 DB 241 GCAAGATTTGCGCGGAATCTCTCAACGAGAGAGTGTATTTCACTGATGAATCTCTT 300

QY 299 GATATATGTAAACCTGTTGACTATATTAACATCTTACCTTCGCTCTAGGCAACC 358
 DB 301 GATATATGTAAACCTGTTGACTATATTAACATCTTACCTTCGCTCTAGGCAACC 360
 QY 359 ATCGTTTCTCCAGATGCTTGAACATTAACAAATTTGGAGAAAGCGCAAAAGAAAGTCAAG 418
 DB 361 ATCGTTTCTCCAGATGCTTGAACATTAACAAATTTGGAGAAAGCGCAAAAGAAAGTCAAG 420
 QY 419 ATCTACTGGGATGTATGATTTTCACTATATAGGATTTATTAACATTAACAGAAACTGA 478
 DB 421 ATCTACTGGGATGTATGATTTTCACTATATAGGATTTATTAACATTAACAGAAACTGA 480
 QY 479 AGTTAGGGAGATGTTGTTCTGCAATTTGCTCTATGCTATGAGGATGTTCAAGGGGCT 538
 DB 481 AGTTAGGGAGATGTTGTTCTGCAATTTGCTCTATGCTATGAGGATGTTCAAGGGGCT 540
 QY 539 GCAATTTCAATTTGAATTCATCATGATTTATTTGAATTTGAGTTCACACTTTTGAGAGA 598
 DB 541 GCAATTTCAATTTGAATTCATCATGATTTATTTGAATTTGAGTTCACACTTTTGAGAGA 600
 QY 599 ATACAGACAGTTATGTTCTGTAATAACTTAATTCCTCATATTTGGAGAAAGAAAG 658
 DB 601 ATACAGACAGTTATGTTCTGTAATAACTTAATTCCTCATATTTGGAGAAAGAAAG 660
 QY 659 TGATGACATTAATTTGAGCCCTTCTCTGCTGCTGAGAAACCTCGTAACGGAGACAAAG 718
 DB 661 TGATGACATTAATTTGAGCCCTTCTCTGCTGCTGAGAAACCTCGTAACGGAGACAAAG 720
 QY 719 AGGTGACAGAAATTAACACAGAGAGACTTAAGATATGCTTTTAACTGATTAACCCAG 778
 DB 721 AGGTGACAGAAATTAACACAGAGAGACTTAAGATATGCTTTTAACTGATTAACCCAG 780
 QY 779 TTTTACTATGTGACAGAAATTAAGATATGCTTACTTATGATGAGAAACCGTGTAG 838
 DB 781 TTTTACTATGTGACAGAAATTAAGATATGCTTACTTATGATGAGAAACCGTGTAG 840
 QY 839 ATATCCCGAGGCAACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACACAG 898
 DB 841 ATATCCCGAGGCAACAGAGCTTGTGAGCAATTTGAGATGACAGCAACATTCACACAG 900
 QY 899 CATAGCCACCTTCTCTGAGCGCTGCTTAAGATTAATTTGACAGGCAAGCTGTGT 958
 DB 901 CATAGCCACCTTCTCTGAGCGCTGCTTAAGATTAATTTGACAGGCAAGCTGTGT 960
 QY 959 CCTGCTACTAGACAGAAAGATTTATCTCTGAGCGATGAGAGGCTAGAAAGCCACTACT 1018
 DB 961 CCTGCTACTAGACAGAAAGATTTATCTCTGAGCGATGAGAGGCTAGAAAGCCACTACT 1020
 QY 1019 TCTTGAAGAACGCAATTTCTATCATTTCTACAGAGTCCAGCCCAATGGGCTTGAGCAACT 1078
 DB 1021 TCTTGAAGAACGCAATTTCTATCATTTCTACAGAGTCCAGCCCAATGGGCTTGAGCAACT 1080
 QY 1081 AATGCTGATCGGATGAGGAGATGAGATGATGATGATGATGATGATGATGATGATGATG 1140
 DB 1139 CCAGATGCTTGAATGACTTTGTGATGTGAATTAAGTGAAGCAATTTATGATCTTTG 1198
 DB 1141 CCAGATGCTTGAATGACTTTGTGATGTGAATTAAGTGAAGCAATTTATGATCTTTG 1200
 QY 1199 GAATCTGTTTGAAGAAACAAAGGTTTATACAGATGCTATCTCTTGGGCAATGTA 1258
 DB 1201 GAATCTGTTTGAAGAAACAAAGGTTTATACAGATGCTATCTCTTGGGCAATGTA 1260
 QY 1259 AGCATTTTCAAGATTTTACGAGAAAGTTCACCGTTATCTCATCTCTTGTGTG 1318
 DB 1261 AGCATTTTCAAGATTTTACGAGAAAGTTCACCGTTATCTCATCTCTTGTGTG 1320
 QY 1319 GAGATTTGTTTGAATTAACATTAAGACCATGAGACTTGTGACTGACCATCAACAA 1378
 DB 1321 GAGATTTGTTTGAATTAACATTAAGACCATGAGACTTGTGACTGACCATCAACAA 1380
 QY 1379 CTGCAATACATCTCTGAGAAATTTGCCGTATATGCTCAGACACACACCAACCAACAA 1438

```

Db      1381  CTGCAATACCACTCCGAGAAATGCGGTAATACCGAG-----TCACATACAGACAGAA 1434
Oy      1439  CAACAGTGTGATCTGTCGCCAGTACTCAACAGACCAACAAATTAACATTTGGATCATCC 1498
Db      1435  CAACAGTGTGATCTGTCGCCAGTACTCAACAGACCAACAAATTAACATTTGGATCATCC 1494
Oy      1499  CAATGACATTAACAAACAAATTAATGTTGACACAGACCAATTAACAGACAGACAAAGT 1558
Db      1495  GAATGACATTAACAAACAAATTAATGTTGACACAGACCAATTAACAGACAGACAAAGT 1553
Oy      1559  AATTAATAGGAAATCTCCGCTTTATGATCCGATTTATGATGATTTGATGATTTATCT 1618
Db      1554  AATTAATAGGAAACACTCCGCTTTATGATCCGATTTATGATGATTTGATGATTTATCT 1613
Oy      1619  TCTTCTCT-AAAAATGTTAGAGCAACAAATTTTATATGTTAGTGTATTTCACT 1677
Db      1614  TCTTCTTAAAAAATTTTATGAGCAACAAATTTTATATGTTAGTGTATTTCACT 1673
Oy      1678  GATTACATTTTATGTTAAAAAATTT 1705
Db      1674  GATTACATTTTATGTTAAAAAATGAT 1701

RESULT 5
US-09-890-475-58
; Sequence 58, Application US/09890475
; GENERAL INFORMATION:
; APPLICANT: Johanson, Urban
; APPLICANT: West, Joanne
; APPLICANT: West, Joanne
; TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late floweri
; FILE REFERENCE: 2001-11-13
; CURRENT FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: PCT/GB00/00197
; PRIOR FILING DATE: 2000-11-25
; PRIOR APPLICATION NUMBER: 98 9902660.1
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 58
; LENGTH: 1715
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
; US-09-890-475-58

Query Match
Best Local Similarity 91.1%; Score 1569.6; DB 33; Length 1715;
Matches 1649; Conservative 0; Mismatches 49; Indels 10; Gaps 4;

Oy      1  CAAGCTTCTCAATTTGGTTG--CTCTCTTACACAGCCCAATCGGTGTTTCGACGT 58
Db      1  CAAGCTTCTCAATTTGGTTGTTCTCTCTTACACAGCCCAATCGGTGTTTCGACGT 60
Oy      59  TTCAGGCTCAATCCAGACATCTATATTAACATATGCAAGAGCGGCTTCAATG 118
Db      61  TTCAGGCTCAATCCAGACATCTATATTAACATATGCAAGAGCGGCTTCAATG 120
Oy      119  TTGCATGAGTTATGCTATACGATAGGAAATTTCAATTTAGGGAGGCTCAGAGTT 178
Db      121  TTGCATGAGTTATGCTATACGATAGGAAATTTCAATTTAGGGAGGCTCAGAGTT 180
Oy      179  TGCATTAATCTCAATTCGGCTCTGACGTTGTTGATGATTTGAACAAGATGTGAG 238
Db      181  TGCATTAATCTCAATTCGGCTCTGACGTTGTTGATGATTTGAACAAGATGTGAG 240
Oy      239  GCAGAAATGTGCGGCAATCTCCACGAGAGAGATTTCAACTGATGAGATCTCT 298
Db      241  GCAGAAATGTGCGGCAATCTCCACGAGAGAGATTTCAACTGATGAGATCTCT 300

Oy      299  GATATATGTAAACCTGTTGACATATTAACATCTTTACCTTGCTCTGTAGCAACC 358
Db      301  GATATATGTAAACCTGTTGACATATTAACATCTTTACCTTGCTCTGTAGCAACC 360
Oy      359  ATCGTTCTTCCAAAGATGCTGATACATTAATTTGAGCAAGACGCAAAAGAAATCAAG 418
Db      361  ATCGTTCTTCCAAAGATGCTGATACATTAATTTGAGCAAGACGCAAAAGAAATCAAG 420
Oy      419  ATCTACGAGATGATGTTTCACTATTAAGATTTGAATTAACATTTACAGAAATCGA 478
Db      421  ATCTACGAGATGATGTTTCACTATTAAGATTTGAATTAACATTTACAGAAATCGA 480
Oy      479  AGTTAGGAGATGATGTTTCACTATTTGCTATGATTTGATTTGATTTGATTTGATTT 538
Db      481  AGTTAGGAGATGATGTTTCACTATTTGCTATGATTTGATTTGATTTGATTTGATTT 540
Oy      539  GCATTTCAATTTGAATTCATCTCATGATTTATTTGATTTGATTTGATTTGATTTG 598
Db      541  GCATTTCAATTTGAATTCATCTCATGATTTATTTGATTTGATTTGATTTGATTTG 600
Oy      599  ATACCAAGACATTAATTTCTGTAAACCTTAATTCCTCATATTTGAGAAAGAGAG 658
Db      601  ATACCAAGACATTAATTTCTGTAAACCTTAATTCCTCATATTTGAGAAAGAGAG 660
Oy      659  TGATGACATTAATTTGAGCCCTGCTCTGCTGCTGCAAACTGTAAGCGAGAGAG 718
Db      661  TGATGACATTAATTTGAGCCCTGCTCTGCTGCTGCAAACTGTAAGCGAGAGAG 720
Oy      719  AGTGGGCAAAATTAACCCAGGAGACTTAAGATGCTTTTACCGTTGATTCACCCAG 778
Db      721  AGTGGGCAAAATTAACCCAGGAGACTTAAGATGCTTTTACCGTTGATTCACCCAG 780
Oy      779  TTTAATTAATGAGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 838
Db      781  TTTAATTAATGAGCAAGAAATTAATTAATTAATTAATTAATTAATTAATTAATTA 840
Oy      839  ATATCCGAGGCAAGAGCTTGTGAGCAATTTGATTAATTAATTAATTAATTAATTA 898
Db      841  ATATCCGAGGCAAGAGCTTGTGAGCAATTTGATTAATTAATTAATTAATTAATTA 900
Oy      899  CATAGCCCACTCTCTGAGAGCGGTGCTTAATTTGATTAATTAATTAATTAATTA 958
Db      901  CATAGCCCACTCTCTGAGAGCGGTGCTTAATTTGATTAATTAATTAATTAATTA 960
Oy      959  CCTGCTACTAAGACAAAGATTAATTCGTCGAGAGCGATTAATTAATTAATTAATTA 1018
Db      961  CCTGCTACTAAGACAAAGATTAATTCGTCGAGAGCGATTAATTAATTAATTAATTA 1020
Oy      1019  TCTTCGAAAGCCCAATTTCTATCTATCTACAGAGTCGCAAGAGCGCTTGAAGAT 1078
Db      1021  TCTTCGAAAGCCCAATTTCTATCTATCTACAGAGTCGCAAGAGCGCTTGAAGAT 1080
Oy      1079  AATGTCGAGCGGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1138
Db      1081  AATGTCGAGCGGATGAGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1140
Oy      1139  CCAGATGCTTATGATCTTTGATGATGATTAATTAATTAATTAATTAATTAATTA 1198
Db      1141  CCAGATGCTTATGATCTTTGATGATGATTAATTAATTAATTAATTAATTAATTA 1200
Oy      1199  GATCTGTTTGAAGAAAGAGGTTATATGATGATGATGATGATGATGATGATGATG 1258
Db      1201  GATCTGTTTGAAGAAAGAGGTTATATGATGATGATGATGATGATGATGATGATG 1260
Oy      1259  AGCATTTTCAAGATTTTACAGAAAGATGATGATGATGATGATGATGATGATGATG 1318
Db      1261  AGCATTTTCAAGATTTTACAGAAAGATGATGATGATGATGATGATGATGATGATG 1320
Oy      1319  GAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1378
Db      1321  GAGATTTTATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1380

```

QY 1379 CTGCAATACATCTCGAGAAATGGCGTAATAGCTACAGACACCACCACCAACAA 1438
 |||||
 Db 1381 CTGCAATACATCTCGAGAAATGGCGTAATAGCTACAGACACCACCACCAACAA 1434
 QY 1439 CAACAGTGTGATGCTCCAGTACTCAACACCAACAAATTAATGATGATTC 1498
 |||||
 Db 1435 CAACAGTGTGATGCTCCAGTACTCAACACCAACAAATTAATGATGATTC 1494
 QY 1499 CAATACATTAACCAACAAAGAAATGAGACAAACAAATTAATGATGATTC 1558
 |||||
 Db 1495 GATATACATTAACCAACAAAGAAATGAGACAAACAAATTAATGATGATTC 1553
 QY 1559 AATTAATAGGAAATCTCCGCTTTATGATACCGATTTATGATGATTC 1618
 |||||
 Db 1554 AATTAATAGGAAATCTCCGCTTTATGATACCGATTTATGATGATTC 1613
 QY 1619 TCTTTCTT-AAAAATGTTAGAGCAACAAATTTTATATGATGATTC 1677
 |||||
 Db 1614 TCTTTCTTAAAAAATGTTAGAGCAACAAAGATTTTATGATGATTC 1673
 QY 1678 GATTAATTTTATGATTAACAAATTAAT 1705
 |||||
 Db 1674 GATTAATTTTATGATTAACAAATTAAT 1701

RESULT 6

US-09-890-220-7
 ; Sequence 7, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; PRIOR FILING DATE: 1999-01-28
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 7
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-7

Query Match 89.3%; Score 1537.6; DB 33; Length 1737;
 Best Local Similarity 95.3%; Pred. No. 0;
 Matches 1649; Conservative 0; Mismatches 49; Indels 32; Gaps 5;

QY 1 CAAGCTTCTCAATTTGCTTG--CTCTCTTACACAGCCAAATCGGTTTTCGACGT 58
 |||||
 Db 1 CAAGCTTCTCAATTTGCTTGCTCTCTCTTACAGCCAAATCGGTTTTCGACGT 60
 QY 59 TTCAGGCTCAATCCAAACAAATCTATATTAAGCAATTTGCAAGAGCGGTTCAATTG 118
 |||||
 Db 61 TTCAGGCTCAATCCAAACAAATCTATATTAAGCAATTTGCAAGAGCGGTTCAATTG 120
 QY 119 TTGCAATGATTTATGCTATGACGTAGAGAAATTTATTTAGGAGGCTCGAGATT 178
 |||||
 Db 121 TTGCAATGATTTATGACGTAGAGAAATTTATTTAGGAGGCTCGAGATT 180
 QY 179 TGCATTAATCTCAATATGCGCTCTTGACGTGTTGATGATTAATTAAGCAAGATGTAG 238
 |||||
 Db 181 TGCATTAATCTCAATATGCGCTCTTGACGTGTTGATGATTAATTAAGCAAGATGTAG 240
 QY 239 GCAGAAATGTCGCGGGAATCTCTACCGGAGAGATGATTTCAACTGATGAAATCTCT 298
 |||||
 Db 241 GCAGAAATGTCGCGGGAATCTCTACCGGAGAGATGATTTCAACTGATGAAATCTCT 300
 QY 299 GATATATTTGAACCTGTTGACATATTAACATCTTTACCTTCGCTCTAGCAACCC 358

Db 301 GATATATTTGAACCTGTTGACATATTAACATCTTTACCTTCGCTCTAGCAACCC 360
 QY 359 ATCGTTTCTCCAGATGCTTTGACTACAAAATTTGAGCAAGAGCCAAAAGATCAG 418
 |||||
 Db 361 ATCGTTTCTCCAGATGCTTTGACTACAAAATTTGAGCAAGAGCCAAAAGATCAG 420
 QY 419 ATCTAGTGGATGATGATTTCACTATTAAGATTTGATTAACATTAACAAAACCTA 478
 |||||
 Db 421 ATCTAGTGGATGATGATTTCACTATTAAGATTTGATTAACATTAACAAAACCTA 480
 QY 479 AGTTAGGAGATGATGATTTGATGATTTGATGATTTGATGATTTGATGATTTGAT 531
 |||||
 Db 481 AGTTAGGAGATGATGATTTGATGATTTGATGATTTGATGATTTGATGATTTGAT 540
 QY 532 -----AGGCGCTCAATTTCAATTTGATGATTTGATGATTTGATGATTTGAT 576
 |||||
 Db 541 CAATATTAACAACTAGGCGGCTCAATTTCAATTTGATGATTTGATGATTTGAT 600
 QY 577 TTGATTTCAAGCTTTTGCAGAAATACAGAGATTAATGTTCTGTAAACCTAATTCCT 636
 |||||
 Db 601 TTGATTTCAAGCTTTTGCAGAAATACAGAGATTAATGTTCTGTAAACCTAATTCCT 660
 QY 637 TCATATTTGAGAGAAAGAGATGATGATGATTAATTTGAGCCCTCTCTCTCTGCA 696
 |||||
 Db 661 TCATATTTGAGAGAAAGAGATGATGATGATTAATTTGAGCCCTCTCTCTCTGCA 720
 QY 697 AACCTGTAGCGGAGACAAAGAGTGGCAGAAATTAACACAGAGACTTAAGATGCT 756
 |||||
 Db 721 AACCTGTAGCGGAGACAAAGAGTGGCAGAAATTAACACAGAGACTTAAGATGCT 780
 QY 757 TTTTACCGTGTGATTCACCCAGTTTACATTAAGAGCAAGAAATGAAATCCACCTA 816
 |||||
 Db 781 TTTTACCGTGTGATTCACCCAGTTTACATTAAGAGCAAGAAATGAAATCCACCTA 840
 QY 817 ATGATGAAACCGTGTGATTTAGATATCCGAGGCAACAGAGCTCTGACAAATTTGAGA 876
 |||||
 Db 841 ATGATGAAACCGTGTGATTTAGATATCCGAGGCAACAGAGCTCTGACAAATTTGAGA 900
 QY 877 TGACCAACAACTTCCACCAAGCATAGCCACTCTCTGTGAGCGCTGTGATAAGTTA 936
 |||||
 Db 901 TGACCAACAACTTCCACCAAGCATAGCCACTCTCTGTGAGCGCTGTGATAAGTTA 960
 QY 937 TATTAACAAAGCAAGCTGTGCTCCGCTGCTACATTAACAAAGATTAATGCTGAGCAT 996
 |||||
 Db 961 TATTAACAAAGCAAGCTGTGCTCCGCTGCTACATTAACAAAGATTAATGCTGAGCAT 1020
 QY 997 CAGAGCTAGAAAGCCACTACTTCTTCAAGAAAGCCAAATTTATATTTCTCAGAGATCC 1056
 |||||
 Db 1021 CAGAGCTAGAAAGCCACTACTTCTTCAAGAAAGCCAAATTTATATTTCTCAGAGATCC 1080
 QY 1057 AGCCATAGGCGCTTGAAGCAATGATGCTGACCGGATGAGAGATGATGATGATGACG 1116
 |||||
 Db 1081 AGCCATAGGCGCTTGAAGCAATGATGCTGACCGGATGAGAGATGATGATGATGACG 1140
 QY 1117 ATGTGCAATTTTAAAGATGCGCAGATGCTGATGATGATGATGATGATGATGATG 1176
 |||||
 Db 1141 ATGTGCAATTTTAAAGATGCGCAGATGCTGATGATGATGATGATGATGATGATGATG 1200
 QY 1177 AAAAGCAATTTCAATGCTTTTGAACCTGTTTGAAGAAACAAAGGGTTATAGCAGATG 1236
 |||||
 Db 1201 AAAAGCAATTTCAATGCTTTTGAACCTGTTTGAAGAAACAAAGGGTTATAGCAGATG 1260
 QY 1237 GTCATATGCTCTTGGCAGATGATGATGATTTTCAAGATTTTGAAGAAAGATGACCTG 1296
 |||||
 Db 1261 GTCATATGCTCTTGGCAGATGATGATGATTTTCAAGATTTTGAAGAAAGATGACCTG 1320
 QY 1297 ACTCATCTCTCTGCTGCTGATGATGATTTTGAATTAATTAAGCAATGATGATGATG 1356
 |||||
 Db 1321 ACTCATCTCTCTGCTGCTGATGATGATTTTGAATTAATTAAGCAATGATGATGATGATG 1380
 QY 1357 TCGACTCAGCAGCAATCAACAACTCAATATGATGATGATGATGATGATGATGATGATG 1416
 |||||

Db 1381 TCGACTCAGCCACATCAACAACTGCAATACATCTCGAATATGCGGTAATACCTCAG 1440
 Oy 1417 ACACGACGACGACCAACAAACAAACAGTGTGATGCTCCAGTGCATCAACACCAACA 1476
 Db 1441 -----TCACCTAACCAACAAACAAAGTGTGATGCTCCAGTGCATCAACACCAACA 1494
 Oy 1477 ACAATACATGTTGGATCANTCCATGACATTAACAACAAGCAATGTGACAGAGG 1536
 Db 1495 ACAATACATGTTGGATCANTCCATGACATTAACAACAAGCAATGTGACAGAGG 1554
 Oy 1537 ACAATACAGACAGACAAATATTAATAGAAAATGCGGGCTTTATGATACGAT 1596
 Db 1555 ACAATACAGACAGAC -AGATATTAATAGAAAACCTCGGTTATGATATCCGAT 1613
 Oy 1597 TTATGGAATGTAATCTTCTCTCTCT -AAAAATGTTAGAGCAAAATTTT 1655
 Db 1614 CTATCGAATGTAATCTTCTCTCTCTTCTTAAAAAATGTTAGAGCAAAATTTT 1673
 Oy 1656 TTTATGTTAGTGTTCATGATGATTTGATTTTAAATTTAAATTT 1705
 Db 1674 TTTATGTTAGTGTTCATGATGATTTTAAATTTTAAATTTAGTTAAATTTGAT 1723

RESULT 7

US-09-513-996A-69038
 ; Sequence 69038, Application US/09513996A
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; FILE REFERENCE: 2750-709P
 ; CURRENT APPLICATION NUMBER: US/09/513, 996A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 81028
 ; SEQ ID NO 69038
 ; LENGTH: 1497
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 1..1497
 ; OTHER INFORMATION: any n or xaa = unknown
 ; FEATURE:
 ; OTHER INFORMATION: Location 1..1497 / Ceres Seq. ID 2185972
 ; US-09-513-996A-69038

Query Match 56.9%; Score 979.8; DB 19; Length 1497;
 Best Local Similarity 89.7%; Pred. No. 1.1e-218;
 Matches 1118; Conservative 0; Mismatches 27; Indels 102; Gaps 2;

Oy 355 ACCCATCGTTCTTCCAGAGATGCTTGAATACAAATTTGAGCAAGGCGCAAAAGAAAGT 414
 Db 320 ACCCATCGTTCTTCCAGAGATGCTTGAATACAAATTTGAGCAAGGCGCAAAAGAAAGT 379
 Oy 415 CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTTGATAATACATTAAGAGAAA 474
 Db 380 CAAGATCTACTGGGATGATGTTTCAACTATTAAGATTTGATAATACATTAAGAGAAA 439
 Oy 475 CTGAAGTTAGGAGAGATGTTCTGTCATATTTGCTCTATGCTATGCTGATGCTTCAAG 534
 Db 440 CTGAAGTTAGGAGAGATGTTCTGTCATATTTGCTCTATGCTGATGCTTCAAG 499
 Oy 535 GGCTGCAATTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTTTTCG 594
 Db 500 GGCTGCAATTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGTACATCT 559
 Oy 595 AAGAAATACAGACAGTTAATGTTCTGTAATACTTAATCTTCATATTTGAGAGAGAG 654
 Db 560 CTGTAATTCACCAACATA----- 577
 Oy 655 GAAGTATGACGATAAATTTGAGCCCTCTCTCTGCTGCAAACTGTAAGGAGAGAC 714
 Db 578 -----TGTTTGCAGCTCGAAACCTGTAAGGAGAGAC 610

Oy 715 AAAGAGTGGCAGAAAATAACACGAGAGACTTAAGTATGCTTTTACCTGTGATTCAC 774
 Db 611 AAAGAGTGGCAGAAAATAACACGAGAGACTTAAGTATGCTTTTACCTGTGATTCAC 670
 Oy 775 CCAGTTTAACATAGCAGAAAATGATATACCTCTTAATGATGAGAAACCTGCTT 834
 Db 671 CCAGTTTAACATAGCAGAAAATGATATACCTCTTAATGATGAGAAACCTGCTT 730
 Oy 835 TAGATATATCCGAGGACAGAGAGCTTGTGAGCAATTTGATGATGACGAGCAATTCAC 894
 Db 731 TAGATATATCCGAGGACAGAGAGCTTGTGAGCAATTTGATGATGACGAGCAATTCAC 790
 Oy 895 CAGGCAATGCGCAGCTCTCTGAGCGCTGCTTAAAGTATATTAACAGAGAGAGCTG 954
 Db 791 CAGGCAATGCGCAGCTCTCTGAGCGCTGCTTAAAGTATATTAACAGAGAGAGCTG 850
 Oy 955 TGTGCTCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGA----- 1007
 Db 851 TGTGCTCTGCTACTAAGACAGAAAGTTATCTGCTGAGCGATCAGAGGCTAGA----- 1007
 Oy 1008 -----AGCCACTACTCTTCTTCAAGAAAGCCAAATTTCTATC 1041
 Db 911 CCATATTTAACGAGAAAATTTCTTCCAGCCACTTCTTCAAGAAAGCCAAATTTCTATC 970
 Oy 1042 ATTCTCAGAGATCCAGCCAAATGCGCTTGAAGCAAGTAAATCTGACCGGATAGGAGG 1101
 Db 971 ATTCTCAGAGATCCAGCCAAATGCGCTTGAAGCAAGTAAATCTGACCGGATAGGAGG 1030
 Oy 1102 ATGAAGTCGATGACAGATGTTGAGATTTGAAGATGCGCAGATGCTTGAATCTTGTG 1161
 Db 1031 ATGAAGTCGATGACAGATGTTGAGATTTGAAGATGCGCAGATGCTTGAATCTTGTG 1090
 Oy 1162 ATGGAATTAAGATGAAGAAAGCAATTCATGATGATGCTTGAATGCTTGAAGAAACAA 1221
 Db 1091 ATGGAATTAAGATGAAGAAAGCAATTCATGATGATGCTTGAATGCTTGAAGAAACAA 1150
 Oy 1222 GGGTTATGACAGATGCTATCTCTGCGCATGGAAGCAATTTCAAGATTTTACGAGA 1281
 Db 1151 GGGTTATGACAGATGCTATCTCTGCGCATGGAAGCAATTTCAAGATTTTACGAGA 1210
 Oy 1282 AAGAGTTGACGCTTACTCATCTCTTCTGCTGCTGAGAGATGTTTGAATTAACATAT 1341
 Db 1211 AAGAGTTGACGCTTACTCATCTCTTCTGCTGCTGAGAGATGTTTGAATTAACATAT 1270
 Oy 1342 GGAACATGAGCTTGTGAGCTGAGCTGAGCCATCAACATGCAATATACCATCTGAGAAAT 1401
 Db 1271 GGAACATGAGCTTGTGAGCTGAGCTGAGCCATCAACATGCAATATACCATCTGAGAAAT 1330
 Oy 1402 GCGGTATGATCTGAGACACCAACACCAACAAACAGAGTGTGATGCTGCCAGTG 1461
 Db 1331 GCGGTATGATCTGAGACACCAACCAACCAACAAACAGAGTGTGATGCTGCCAGTG 1390
 Oy 1462 ACTCAACACCAACCAACATTAACATTTGATGATCCATGACATTAACACAGAGACA 1521
 Db 1391 ACTCAACACCAACCAACATTAACATTTGATGATCCATGACATTAACACAGAGACA 1450
 Oy 1522 ATGTGCAACAAGAGCAATTAACAGAGAGCAAGTATTAATATG 1568
 Db 1451 ATGTGCAACAAGAGCAATTAACAGAGAGCAAGTATTAATATG 1497

RESULT 8

US-09-890-220-3
 ; Sequence 3, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Genbail, Anthony
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Newburt
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27

PRIOR APPLICATION NUMBER: PCT/GB00/00248
 PRIOR FILING DATE: 2000-01-28
 PRIOR APPLICATION NUMBER: GB 9901927.5
 PRIOR FILING DATE: 1999-01-28
 NUMBER OF SEQ ID NOS: 77
 SOFTWARE: Patent Ver. 2.1
 SEQ ID NO 3
 LENGTH: 6338
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-890-220-3

Query Match 23.9%; Score 412.2; DB 33; Length 6338;

Best Local Similarity 98.8%; Pred. No. 2,1e-85;

Matches 414; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1304 ACTCTTCTGCTGGAGATTTGTTTGAATTAACCTAGGAACCATGACCTTGCTGACTC 1363
 DB 5102 AATTTCCAGCTGTGGAGATTTGTTTGAATTAACCTAGGAACCATGACCTTGCTGACTC 5161
 QY 1364 ACCACGATCAACACTGCAATACATCTCTCGAATTTCCGTAATAGCTGACACCCAC 1423
 DB 5162 ACCACGATCAACACTGCAATACATCTCTCGAATTTCCGTAATAGCTGACACCCAC 5221
 QY 1424 CACCAACAACAACAAGTGTGATCGTCCAGTGCATCAACACCAACAACAATTA 1483
 DB 5222 CACCAACAACAACAAGTGTGATCGTCCAGTGCATCAACACCAACAACAATTA 5281
 QY 1484 CATTTGGATCTCCCAATGACATTAACAACAACAAGTGTGACACACAGACATTA 1543
 DB 5282 CATTTGGATCTCCCAATGACATTAACAACAACAAGTGTGACACACAGACATTA 5341
 QY 1544 CAGCAGACCAAGTAATTAATAGAAATCTCGGCTTTATGATACGATTTATCGG 1603
 DB 5342 CAGCAGACCAAGTAATTAATAGAAATCTCGGCTTTATGATACGATTTATCGG 5401
 QY 1604 ATTTACTATTTCTCTTCTTAAATAATGTTAGAGCAACAATTTTATATGT 1663
 DB 5402 ATTTACTATTTCTCTTCTTAAATAATGTTAGAGCAACAATTTTATATGT 5461
 QY 1664 TAGGTATTTCACTGATTTATTTAGTTAAAAAAAATGATTTGCTTATATCT 1722
 DB 5462 TAGGTATTTCACTGATTTATTTAGTTAAAAAAAATGATTTGCTTATATCT 5520

RESULT 9

US-10-219-999-5373

Sequence 5373, Application US/10219999

GENERAL INFORMATION:

APPLICANT: Cao, Yongwei

APPLICANT: Edgerton, Michael D

APPLICANT: Hinkle, Gregory J.

APPLICANT: Kovalic, David K.

APPLICANT: Liu, Jingdong

APPLICANT: Stein, Joshua

TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT

FILE REFERENCE: 38-10(52726)C

CURRENT APPLICATION NUMBER: US/10/219,999

PRIOR FILING DATE: 2002-08-15

PRIOR APPLICATION NUMBER: US 60/324,109

PRIOR FILING DATE: 2001-09-21

PRIOR APPLICATION NUMBER: US 60/312,544

PRIOR FILING DATE: 2001-08-15

NUMBER OF SEQ ID NOS: 63520

SEQ ID NO 5373

LENGTH: 1722

TYPE: DNA

ORGANISM: glycine max

FEATURE:

NAME/KEY: CDS

LOCATION: (164)..(1450)

OTHER INFORMATION:

US-10-219-999-5373

Query Match 21.3%; Score 366.8; DB 42; Length 1722;
 Best Local Similarity 59.8%; Pred. No. 5.7e-75;
 Matches 757; Conservative 0; Mismatches 422; Indels 87; Gaps 5;

QY 228 AGAATGTAGCAGAAATTTGCGCGGAATCTCCACCGAGAGATGATTTCACTGAT 287
 DB 161 AGCATGTCCGGGAAATTTCTCCGGTACACATCCCGGTGAAGAAATAATTTGACGTGAT 220
 QY 288 GAGATCTCTGATATATTTGTAACCTGTTCGACTATATATACATCTTTCGCTCT 347
 DB 221 GAGAGCTTTGATTTATTTGCAAGCTGTGAACTTTACAAATTTCTACCCGCTGCT 280
 QY 348 CTAGCAACCCATCTTTCTTCCAAATGCTTGAATTAATAAATTTGACCAAAAGCCAAA 407
 DB 281 CTTCAAAATCTTCTTCTTTAGAGATTTTGGCTTAATAAATTAAGACCAAGCCGTTAA 340
 QY 408 AGAAATCAAGATCTACTGGATGTGATTTTCACTATATAGATTTGATTAACATTA 467
 DB 341 AGGAGTTGAG---GCAGAAATTTGATTTCAATTTATAGGACCGCTACCAATTTCT 397
 QY 468 CAGAAATCTGAAGTTAGGAGATTTGCTTGTCCATTTTGTCTATATGCTATGTGTAC 527
 DB 398 CGAAAGCTGAAGTGAACCAAGACTTTTGTGCTTTTGTGCTGATGAGGTGGAAC 457
 QY 528 TTCAAGGGCTGCAATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 587
 DB 458 TTTAAGGTTTGGATTTATCTTTTGTATTCATGATTTGATTTGATTTGATTTGATTT 517
 QY 588 CTTTTCGAAGATTAACAGACAGTAAATGTTTCTGTAATAATCTTCAATTTCTCATATTTGAG 647
 DB 518 GTTACTGAAGATTTACCAACAGTGAATGATTTCTGTAATAATGATTTATGATGATGAG 577
 QY 648 GAAGAAGAGATGATACAT---AAATTTGAGCCCTTCTCTCTGCTGAAACCTGCT 704
 DB 578 AATGTTGCTGATGAGATTAATTCACAAATGCAAACTTCTTCTGCTGCAAGACCTGCA 637
 QY 705 AAGCGGAGCAAAAGAGTGGAGAAATTAACACAGAGACTTAAATGATTTGATTTGATTT 764
 DB 638 AAACGTAGAAAGAAAGACTCTGTTCAAAATTTGAAAAACGCAATTAATTTCTGAGAG 697
 QY 765 TTGGATTCACCA----- 777
 DB 698 TTGAGACTCACAGAAAGCATATATATGATTTCTCAAAAGATGATATCTATATCC 757
 QY 778 -----GTTAATCTAATGACACAGAAATGGAATACCTCTACTTATATGA 820
 DB 758 TCCAAGAGAGAGATGTTGCTAGAACATCTGTAAGTGAAGAAATTTTCCCTAGTGAAGA 817
 QY 821 T-----GGAACCGTGGTTTAGATATCCGAGCAACAGAGCTTCTGACAAATT 872
 DB 818 AATGACGAGGAAATTTTGTCTGATCATCTCGGACCATGACAACTGAGGACATGTG 877
 QY 873 GAGATGACCAAGCAATTCACCA-----GCCATAGCCACTCTTCTGACGCT 923
 DB 878 GATCCAGTTTCAACATCTCAGGTGTTCAATTTGCAATGCCCAATCTTCTGAGCCCT 937
 QY 924 GGTGCTAAAGTTATATTTGACAAAGCAAGCTGTGCTCTGCTTACTTAAGCAAGAAAGTTA 983
 DB 938 GATGTAGTAAATCATATATATTAAGTATCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 997
 QY 984 TCTGCTGAGCTGACAGAGCTGAGAGCAACCTACTCTTCAAGAAAGCCAAATCTATAT 1043
 DB 998 AGCATGATCTGATCAGATCTCAGAAACCAATCTCTGCTGACAGAGAGATTTCTTTTCC 1057
 QY 1044 TCTCAGAGCTCAGGCAATTTGAGCAAGATATGCTGACCGGATAGAGAGAT 1103
 DB 1058 TCACACAGAGTCCAGCTATGACACTAGAAACAAGTATATCAGACGCTATAGTGAAGAC 1117
 QY 1104 GAAGTCGATGACAGATTTGACATTTTGAAGATCCGCAATGCTGATATGATCTTGTGAT 1163
 DB 1118 GAAGTTGATGACGACATTTGACATTTTGAAGATGAAGAGATGCTTGAAGATTTTGTGAT 1177

QY 1164 GTGATTAAGATGAAAGCAATTCATGCTTTGGAAGTCTTTGTAAGAAACAAAGG 1223
DB 1178 GTTCCAAAGATGAAAGCAATTCATGCTTTGGAAGTCTTTGTAAGAAACAAAGG 1237
QY 1224 GTTATAGCATGATGATATCTCTTGGGATGAGCAATTTTCAAGATTTTACGAGAA 1283
DB 1238 GTGCTAGCATGATGATATCTCTTGGGATGAGCAATTTTCAAGATTTTACGAGAA 1297
QY 1284 GAGTTCACCGCTTACTCATCTCTTGGTGTGGAGATTTTGTATTAACATATG 1343
DB 1298 GAGCTGATCTCATCTCTTGGTGTGGAGATTTTGTATTAACATATG 1357
QY 1344 AACCATGAGCTTGTGCTGACATGACCAATCAACATGCAATCTCGAATTCG 1403
DB 1358 AATCATGCTCTTGTGCTGACATGACCAATCAACATGCAATCTCGAATTCG 1417
QY 1404 CGTAAT 1409
DB 1418 AGAAT 1423

RESULT 10
US-09-565-309A-32201
Sequence 32201, Application US/09565309A
GENERAL INFORMATION:
APPLICANT: ALEXANDROV, Nikolai
TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
FILE REFERENCE: 2750-0853P
CURRENT APPLICATION NUMBER: US/09/565, 309A
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 68449
SEQ ID NO 32201
LENGTH: 606
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)..(606)
OTHER INFORMATION: any n = a, g, c, t, unknown, or other
NAME/KEY: misc-feature
LOCATION: (1)..(606)
OTHER INFORMATION: 39533:110036 (Clone Number: Unique Sequence Identifier)
US-09-565-309A-32201

Query Match 19.0%; Score 326.4; DB 22; Length 606;
Best Local Similarity 98.0%; Pred. No. 1.1e-65;
Matches 341; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
QY 1373 CAACAAGTGCATATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 1432
DB 259 CAACAAGTGCATATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 318
QY 1433 CAACAAGTGCATATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 1492
DB 319 CAACAAGTGCATATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 378
QY 1493 TCATCCCATGATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 1551
DB 379 TCATCCCATGATGATGCTGAGAAATGCGCTTAATATAGCTCAGACCAACCAACCA 438
QY 1552 ACAAGTAATTAATAGGAAATGCGCTTTATATAGCTCAGATTTATCGATTTAATC 1611
DB 439 ACAAGTAATTAATAGGAAATGCGCTTTATATAGCTCAGATTTATCGATTTAATC 498
QY 1612 TATATCTCTCTTCTTAAATTTTATAGGAGCAACAAATTTTATATAGTATGAT 1671
DB 499 TATATCTCTCTTCTTAAATTTTATAGGAGCAACAAATTTTATATAGTATGAT 558
QY 1672 TCACTGATTAATTTTATAGTATGATGATGATGATGATGATGATGATGATGATGAT 1719
DB 559 TCACTGATTAATTTTATAGTATGATGATGATGATGATGATGATGATGATGATGAT 606

RESULT 11
US-09-890-220-6
Sequence 6, Application US/09890220
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
TITLE OF INVENTION: Methods and means for modification of plant characteristics us
FILE REFERENCE: Mewburn
CURRENT APPLICATION NUMBER: US/09/890,220
CURRENT FILING DATE: 2001-07-27
PRIOR APPLICATION NUMBER: PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
PRIOR FILING DATE: 1999-01-28
NUMBER OF SEQ ID NOS: 77
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 6
LENGTH: 5895
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-6

Query Match 18.5%; Score 319.4; DB 33; Length 5895;
Best Local Similarity 92.7%; Pred. No. 1.2e-63;
Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;
QY 1308 TTCTGTTGGAGATGTTTGTATTAATGAAACATGAGACTTGTGACCTGAGCC 1367
DB 4677 TCCAGGTGTTGAGATGTTTGTATTAATGAAACATGAGACTTGTGACCTGAGCC 4736
QY 1368 ACCATCAACATGCAATACATCTCTGAGATTCGCTTAATAGCTGACACCAAC 1427
DB 4737 ACCATCAACATGCAATACATCTCTGAGATTCGCTTAATAGCTGACACCAAC 4790
QY 1428 ACCAACAACATGCAATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 1487
DB 4791 ACCAACAACATGCAATGATGCTGCTGATGCTGATGCTGATGCTGATGCTGATG 4850
QY 1488 GTGATCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 1547
DB 4851 GTGATCATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATG 4910
QY 1548 AGAGACAAGTAATTAATAGGAAATGCGCTTTATATAGCTCAGATTTATCGATTTG 1607
DB 4911 AGAGACAAGTAATTAATAGGAAATGCGCTTTATATAGCTCAGATTTATCGATTTG 4969
QY 1608 TAACTATTTCTTCTTCTT-AAAAATTTGTTAGAGCAACAAATTTTATATAGTATG 1666
DB 4970 TAACTATTTCTTCTTCTT-AAAAATTTGTTAGAGCAACAAATTTTATATAGTATG 5029
QY 1667 TGTATTCACGATTAATGATTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 1705
DB 5030 TGTATTCACGATTAATGATTTTATGTTTATGTTTATGTTTATGTTTATGTTTATG 5068

RESULT 12
US-09-534-859-587
Sequence 587, Application US/09534859
GENERAL INFORMATION:
APPLICANT: Bush, David F.
APPLICANT: Last, Robert L.
APPLICANT: Levin, Irene M.
APPLICANT: Norris, Susan R.
APPLICANT: Parnell, Laurence D.
APPLICANT: Rounsley, Steven D.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
FILE REFERENCE: 38-10(15493)B
CURRENT APPLICATION NUMBER: US/09/534, 859
CURRENT FILING DATE: 2000-03-29

NUMBER OF SEQ ID NOS: 1127
 SEQ ID NO 587
 LENGTH: 201471
 TYPE: DNA
 ORGANISM: Arabidopsis thaliana
 US-09-534-859-587

Query Match 18.5%; Score 319.4; DB 20; Length 201471;
 Best Local Similarity 92.7%; Pred. No. 4.6e-63;
 Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

QY 1308 TTCGTGTTGGAGATTGTTTGAATTAACATATGGAACCATGAGCTGTGCATCAGCC 1367
 DB 42342 TCCAGGTGTGGAGATTGTTTGAATTAACATATGGAACCATGAGCTGTGCATCAGCC 42401
 QY 1368 ACCATCAACAACTGCAATATCCATCTCCGAGAAATTCGGTAATAGCTGACAGCACACAC 1427
 DB 42402 ACCATCAACAACTGCAATATCCATCTCCGAGAAATTCGGTAATAGCTGACAGCACACAC 14255
 QY 1428 ACCAACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATACATT 1487
 DB 42456 AACCAACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATACATT 42515
 QY 1488 GTGATCATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGC 1547
 DB 42516 GTGATCATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGC 42575
 QY 1548 AGAGCAAAAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGATTG 1607
 DB 42576 AGAGCAAAAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGATTG 42634
 QY 1608 TAACTTATCTCTTCTTCTT-AAAAAATGTTAGAGCAACAATAATTTTATATGTAG 1666
 DB 42633 TAACTTATCTCTTCTTCTTAAAAAATGTTAGAGCAACAATAATTTTATATGTAG 42694
 QY 1667 TGTATCAACTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1705
 DB 42695 TGTATCAACTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 42733

RESULT 13

US-09-803-736-587
 ; Sequence 587, Application US/09803736
 ; GENERAL INFORMATION:
 ; APPLICANT: Bush, David F.
 ; APPLICANT: Levin, Irena M.
 ; APPLICANT: Norris, Susan R.
 ; APPLICANT: Roundsley, Steven D.
 ; APPLICANT: Wiegand, Roger C.
 ; TITLE OF INVENTION: Plant Polymorphic Markers and Uses Thereof
 ; FILE REFERENCE: 38-10(15493)D
 ; CURRENT APPLICATION NUMBER: US/09/803,736
 ; CURRENT FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: US 09/534,859
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: Identified by Attorney Docket number 04983.0206CPUS01 38-10
 ; NUMBER OF SEQ ID NOS: 1582
 ; SEQ ID NO 587
 ; LENGTH: 201471
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-803-736-587

Query Match 18.5%; Score 319.4; DB 31; Length 201471;
 Best Local Similarity 92.7%; Pred. No. 4.6e-63;
 Matches 370; Conservative 0; Mismatches 21; Indels 8; Gaps 3;

QY 1308 TTCGTGTTGGAGATTGTTTGAATTAACATATGGAACCATGAGCTGTGCATCAGCC 1367
 DB 42342 TCCAGGTGTGGAGATTGTTTGAATTAACATATGGAACCATGAGCTGTGCATCAGCC 42401
 QY 1368 ACCATCAACAACTGCAATATCCATCTCCGAGAAATTCGGTAATAGCTGACAGCACACAC 1427

DB 42402 ACCATCAACAACTGCAATATCCATCTCCGAGAAATTCGGTAATAGCTGACAGCACACAC 1487
 QY 1428 ACCAACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATACATT 1487
 DB 42456 AACCAACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATACATT 42515
 QY 1488 GTGATCATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGC 1547
 DB 42516 GTGATCATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGC 42575
 QY 1548 AGAGCAAAAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGATTG 1607
 DB 42576 AGAGCAAAAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGATTG 42634
 QY 1608 TAACTTATCTCTTCTTCTT-AAAAAATGTTAGAGCAACAATAATTTTATATGTAG 1666
 DB 42633 TAACTTATCTCTTCTTCTTAAAAAATGTTAGAGCAACAATAATTTTATATGTAG 42694
 QY 1667 TGTATCAACTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1705
 DB 42695 TGTATCAACTGATTAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 42733

RESULT 14

US-09-513-996A-69042
 ; Sequence 69042, Application US/09513996A
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
 ; FILE REFERENCE: 2750-709P
 ; CURRENT APPLICATION NUMBER: US/09/513,996A
 ; CURRENT FILING DATE: 2000-02-25
 ; NUMBER OF SEQ ID NOS: 81028
 ; SEQ ID NO 69042
 ; LENGTH: 319
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: UNSURE
 ; LOCATION: 1..319
 ; OTHER INFORMATION: any n or Xaa = unknown
 ; FEATURE:
 ; OTHER INFORMATION: Location 1..319 / Ceres Seq. ID 2185983
 US-09-513-996A-69042

Query Match 18.2%; Score 314.2; DB 19; Length 319;
 Best Local Similarity 99.1%; Pred. No. 6.3e-63;
 Matches 316; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1375 ACAACTGCAATACATCTCCGAGAAATTCGGCTTAATAGCTGACAGCACACACACCAACA 1434
 DB 1 ACAACTGCAATACATCTCCGAGAAATTCGGCTTAATAGCTGACAGCACACACACCAACA 60
 QY 1435 ACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATGATGATC 1494
 DB 61 ACAACAACAGTGTGATGTCGCCAGTACTCAACACCAACAATATGATGATC 120
 QY 1495 ATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGAGACA 1554
 DB 121 ATCCCAATGACATTAACAACAAGAAATGTTGACACAGGACAAATAACAGAGACA 180
 QY 1555 AAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGGATTAACCTTA 1614
 DB 181 AAGTAATTAATAGGAATAATCCCGCTTTATGATACCGATTATCGGATTAACCTTA 240
 QY 1615 TTCCTCTTCTTAAAAAATGTTAGAGCAACAATAATTTTATATGTAGTATCA 1674
 DB 241 TTCCTCTTCTTAAAAAATGTTAGAGCAACAATAATTTTATATGTAGTATCA 300
 QY 1675 ACTGATTACATTTTATGTT 1693

Fri Jun 20 09:03:39 2003

us-09-890-220-1.rtf

Page 12

```

Db          301 ACTGATTACATTTTAGT 319

RESULT 15
US-60-303-460-9
: Sequence 9, Application US/60303460
: GENERAL INFORMATION:
: APPLICANT: Allen, Steve
: APPLICANT: Butler, Karlene
: APPLICANT: Niu, Xiping
: TITLE OF INVENTION: Flowering Regulating Polynucleotides and Polypeptides
: FILE REFERENCE: B81493 US, PRV
: CURRENT APPLICATION NUMBER: US/60/303,460
: CURRENT FILING DATE: 2001-07-06
: NUMBER OF SEQ. ID NOS: 19
: SOFTWARE: Microsoft Office 97
: SEQ. ID NO. 9
: LENGTH: 2346
: TYPE: DNA
: ORGANISM: Helianthus sp.
: US-60-303-460-9

```

Query Match	18.28;	Score 314;	DB 74;	Length 2346;
Best Local Similarity	58.40;	Pred. No. 1.5e-62;		
Matches 598; Conservative	0;	Mismatches 405;	Indels 21;	Gaps 2

```

OY      1101 GATGAAGTCGATGACGATGTTGTGACGATTTTGAAGATCCGACAGTCGTTGATACCTTGTG  1160
Db      1646 GATGAAGTGTGAGATGATGTTGCTGATCTCGSAGACCGTAGCATGCTGTGGACGACTTCTGT  1705

OY      1161 GATGCGAATAAAGATGAAAGAACCAATTCATTCATCTTTGGAGCTCGTTGTAAAGAAACAA  1220
Db      1706 GATGTTTCCAAAGACGAAACAAACAGATGATGATCATCTATGGAACATCTATTGTGAGAAAGCAA  1765

OY      1221 AGGGTTATATAGCAGATGTCATATCTCTTTGGGCATGTGAAGACATTTTCAAGATTTTACGAG  1280
Db      1766 CCGGTCGTGCGGCGSAGATGGCCATATTTCTGTGCTGTGGAAAGCGTTTCAAAAGCTGCATGCT  1825

OY      1281 AAAAGGTGCAACCGTTACTACTACTACTCTCTGCTGTGGAGATGTTTGTGATTAACATA  1340
Db      1826 CAAGACCTTGTTCAGAGCTCCATCTACTCTCTGCTGTGGAGATGTTTGTGATTAACATA  1885

OY      1341 TGGAAACATGAGACTGTGCACTGACGACACATATAAATGCAATACATCTCTGAGAAAT  1400
Db      1886 TGGAAATCATGGGCTCTTGACTCTCCGAGACCTTAACCAATTGTAAATGTGATCTTGACAA  1945

OY      1401 TGCC 1404
Db      1946 TACC 1949

```

Search completed: June 19, 2003, 23:17:45
Job time : 4034 secs

QY	384	TACAAAATTTGGAGCAAAACGCGAAAAAGAAAGTCAAGATCTCTGGGATGTGAGTATTTTCAAC	443
Db	944	TACAAACGATGCTCTAGACCCGGGTGTATGAGATTGGGGTCTGGAAATGCGTTTTAAT	1003
QY	444	TATTAAGATTGTAAATTAACACATTTACAGAAAACCTAAGTTAGGGAGATGTTCCTGTCCA	503
Db	1004	TACAAATACATAACAAATACATTCGATAGAGACTAATGTATCAAGAGATTTTCTTGTCCG	1063
QY	504	TTTTGGCTATACCTATTTGGTAGCTTCAAGGAGGCTGCATTTATTTGAATTCATCTCAT	563
Db	1064	TTCTGTTTGGTAAATATGCGTAGATTACAAAGGGCTGTGAATTTTCACTTGACTTCATCATAT	1123
QY	564	GATTTTATTTGAATTTAGTTCAAAGCTTTTCCAAATATACAGACAGTATATGTTCTGTGA	623
Db	1124	GATCTCTTCAATTAAGATTTTGGTTTGGTTACTGAACATGTCAGTCCGATGTGATATATAC	1183
QY	624	AAACTTAATTTCTTATATTTTGGAGAA--AAAGAGATGATACATTAATTTTGGGCC	680
Db	1184	AGAACCGATGCAAGACGCTCGAATATTAATTCGTAGAAACCTTATCCGAACAAACAAATG	1243
QY	681	TTTCTCTCTGTGTCAAACCTCGTAAAGGGAGAACAAAGAGTGGGCAAAATATTAACCCAGG	740
Db	1244	TTTCTCTTATTTGGCAATTAATCGAAGAAAGCCGGGAAACAAAACCTCAATCTCAAAACGA	1303
QY	741	ACAGCTTAAGATATGCTTTTACGGTGGATTCACCCGATTTACTAATAGGCAACAGAAAT	800
Db	1304	AATCAT-----GTGATCTCACTGTGCTGGATTTCAACATGTCTTACA	1345
QY	801	GGAAATCACCCATTAATGATGAGAAACCGGTGTATAGATATCCGAGGCAACACAGCTT	860
Db	1346	GCTCTCAGTAGACCTTATAGATTAATACAGATGTGTACTGATGACATAGAGCTGATGTCA	1405
QY	861	GCTGACAAATTTGAAATGATACAGGACAATTCACACAGCCATAGCCACTCTTCTCTGAC	920
Db	1406	TGTATGTCGGATGCAAGTGCACATGTCCCTCAATATGCAACCCGAGGCTTATACATGCA	1465
QY	921	GCTGGTGTCAAAAGTTATTTGACAAAGGAAGCTGTGCTCTCTCTCTCTCTTAAGCAAGAAAG	980
Db	1466	TTTACCTGAAGACGACACCTTGCCACCTTCACCCATGTGTATGTTTCCAAAGACGAGAAAG	1525
QY	981	TTTATCTGCTGAGCCATCAAGAGCTTGAAGCCACCTTACTTCTTACAAACGCCCAATTTTAT	1040
Db	1526	TTTATGCGTTGACCGATCCGATCCGTGAAGAACCGAGCCCTGCTGTGAGAAAGGGAGTTTCTTT	1585
QY	1041	CATTCTCAAGAGTCCAGCCCAATGGCGCTTGAGCAAGTATGTCTATACCGGGATAGCGAG	1100
Db	1586	CACCTCCACCGAGCTCACCCCAATGGCGATGAACAGATTTGCTGTGAGGAACAGTGA	1645

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 19:25:25 ; Search time 2561 Seconds
(without alignments)
10889.742 Million cell updates/sec

Title: US-09-890-220-1

Perfect score: 1722

Sequence: 1 caagctcttcatttctgct.....aatgattctgtctataact 1722

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	628.2	36.5	672	10	AV822548
2	266.6	15.5	526	17	CNS00MYB
3	237.8	13.8	680	10	AM038171
4	232.4	13.5	769	14	B0505017
5	226	13.1	1128	11	AV104964
6	212	12.3	846	12	BG648271

7	206.2	12.0	480	9	AJ469021
8	203	11.8	569	14	B0611550
9	201.4	11.7	600	13	B1479743
10	198.4	11.5	474	10	AM234600
11	197.8	11.5	391	10	AV817525
12	190	11.0	561	14	BQ740672
13	188.8	11.0	467	12	BG580495
14	176.4	10.2	364	9	B1164598
15	174.6	10.1	405	12	BG314044
16	173.6	10.1	554	13	B1784785
17	171.2	9.9	546	13	B1321750
18	170.2	9.9	444	12	BG551103
19	169.6	9.8	612	10	BE203328
20	164	9.5	429	14	BQ459233
21	160.6	9.3	445	10	AM202144
22	154.6	9.0	326	9	A1163743
23	144.2	8.4	561	14	BQ162202
24	143	8.3	662	14	BQ279630
25	132.8	7.7	519	14	BQ168835
26	126	7.3	644	13	B1453006
27	125.6	7.3	322	10	AM309408
28	124	7.2	757	13	BQ460553
29	121	7.0	515	12	BF481866
30	120.6	7.0	236	14	BQ904049
31	113.6	6.6	560	9	A1795561
32	113.6	6.6	590	12	BG649311
33	113.6	6.6	594	10	AM000259
34	110.6	6.4	586	14	BQ243123
35	109.2	6.3	727	14	BQ801572
36	108.6	6.3	732	14	BQ841916
37	105.8	6.1	142	10	AV783345
38	104.8	6.1	463	14	C72616
39	101.2	5.9	576	10	AM076157
40	100.4	5.8	675	13	B1932726
41	94.8	5.5	523	13	B1992698
42	94	5.5	553	12	BG909756
43	94	5.5	553	14	BQ620168
44	94	5.5	593	13	BM377912
45	93.6	5.4	540	10	BE587348

ALIGNMENTS

RESULT 1	AV822548	672 bp	mRNA	linear	EST 01-APR-2002
LOCUS	AV822548	RAFL5	Arabidopsis thaliana	cDNA clone	RAFL05-08-K19 5',
DEFINITION	mRNA sequence.				
ACCESSION	AV822548				
VERSION	AV822548.1	GI:19864597			
KEYWORDS	EST.				
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
REFERENCE	Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis. 1 (bases 1 to 672)				
AUTHORS	Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M., Oono,Y., Sakurai,T., Carninci,P., Kawai,J., Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shingawa,A., Muramatsu,M., Hayashizaki,Y. and Shinozaki,K.				
TITLE	Large scale analysis of Arabidopsis full-length cDNA (2002b)				
JOURNAL	Unpublished (2002)				
COMMENT	Contact: Motoaki Seki Plant Functional Genomics Research Group RIKEN Genomic Sciences Center 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan Tel: 81-298-36-4359 Fax: 81-298-36-9060 Email: mseki@rcl.riken.go.jp An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). This clone is in a				

REFERENCE
D'Ascenzo, M., He, X., Lyman, J., Holt, I.E., Liang, F., Upton, J.,
Roning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nieman, W.,
Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni,
J.

TITLE
Generation of ESTs from tomato leaf tissue

JOURNAL
Unpublished (1999)

COMMENT
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES
source
1.680
/location=/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="Rio Grande Flor"
/db_xref="taxon:4081"
/clone="CLER1N5"
/clone_lib="tomato mixed elicitor, B1"
/tissue_type="leaf"
/dev_stage="4-6 week old plants"
/lab_host="X11-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLET - inoculated with a variety of disease response
elicitors. Plants exposed to 2.6 dichloroisocitric
acid, BTH, jasmonic acid, ethylene, fenthion, EIX,
okadaic acid, or systemin prior to tissue harvest. EcoRI
site was destroyed during cloning."

BASE COUNT
ORIGIN
192 a 120 c 164 g 204 t

Query Match
Best Local Similarity 74.1%; Score 237.8; DB 10; Length 680;
Pred. No. 2.1e-42;
Matches 315; Conservative 0; Mismatches 107; Indels 3; Gaps 1;

987 GGTGACGATCAGAGGTGAGAGCCACCTCTTTCAGAGCCCAATTCATCTCTT 1046
1 GTTGGACGCTGTGATCCCGAAGATGTGCTCTCTGCAAAAAGCAATTCCTTCTT 60
1047 CACAGAGTCCAGCCCAATGCGCTTGAGCAAGTATGTCTGACCGGATAGCGAGTGA 1106
61 CATAGGCCCCAGCCCATGCTGACCTGAGCAAGTTTATCGACGAGACAGTGA 120
1107 GTTCGATGAGATGTTCGAGATTTGAAGATGCCAGATCTGATGACTTTTGTGATG 1166
121 GTTGATGATGATGTTCGAGATCTTGAAGATCGAAGATCCTTGATGATTTTGTGATG 180
1167 AATTAAGATGAAGCAATTCATGATCTTTGGAATCTTTGTAAGAAAAGAGGTT 1226
181 ACCAAGATGAAGCAAGTATGATCTGTGAACTCTTTGTTAGAAAGCAAGG 240
1227 ATAGCAGATGTCATATCTCTTGGCAGTGAAGATTTTCAAGATTTTACAGAAAG 1286
241 TTGGCAGATGTGATATCCCTTGGCAGTGTGAGCCCTTTCAAGTCGATGTGAGAG 300
1287 TTGCACCGTTACTCA---TCACTCTTGTGCTGTGAGATTTTGTGTTAACTATG 1343
301 TTGGCCCAAGACACGACACTATGACAGGTGTGGAGTTATTCATGATGAGGCTG 360
1344 AACCATGACCTGTGACACTCAGACCAATCAACACTGATATACCATCTCCGAAATTC 1403
361 AACCATGCGCTGTGATGCGCGGTACAAATTAACATTTGTAACCTAATATTAAGCAG 420
1404 CGTAA 1408
421 CAAAA 425

RESULT 4
B0505017/c B0505017 769 bp mRNA linear EST 22-JUL-2002

DEFINITION
EST612432 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMGB40
3' end, mRNA sequence.

ACCESSION
B0505017

VERSION
B0505017.2 GI:21920982

KEYWORDS
EST.

SOURCE
Solanum tuberosum
potato.

REFERENCE
Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, M., Smart, C.,
Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and
Karamycheva, S.A.

AUTHORS
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21363886.
Other ESTs: EST612431

COMMENT
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato@ligr.org
This clone is available through the Research Genetics, contact the
Research Genetics for further information 1-800-711-6195 or
cdna@resgen.com
Seq primer: 17.

FEATURES
source
1.769
/location=/Qualifiers
/organism="Solanum tuberosum"
/cultivar="Kennebec or Blajete"
/db_xref="taxon:4113"
/clone="STMGB40"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating eyes
tubers, or roots."

BASE COUNT
ORIGIN
239 a 176 c 140 g 214 t

Query Match
Best Local Similarity 73.0%; Score 232.4; DB 14; Length 769;
Pred. No. 3.4e-41;
Matches 327; Conservative 0; Mismatches 116; Indels 5; Gaps 2;

961 CTGCTACTAGCAAGAAAGTATCTGCTGAGCGATCAGAGGCTGAAGCCACCTACTTC 1020
747 CAGTTCGCCAAGTCAGAAAGTTATTCAGTTGAGGCTCTATCCCGAATTCGTGCTCC 688
1021 TTCAGAAAGCCCAATTCATTCATTCACAGAGTCCAGCCAAATGGCGTTGAGCAAGTAA 1080
687 TCCAAAAAAGGCAA-TCTTTCATTCATCAATGAGGCCAGCCATGCGACTGAGCAAGTT 629
1081 TGTCTGACCGGATAGCAGATGAAGTGAATGATGACAGATTTGCAATTTTGAAGATGCC 1140
628 TATTCAGATCGAACACAGTGAAGTGAAGTGAATGATGATTTTCAGATCTTGAAGATCGAA 569
1141 AGATGCTGATGACTTTGAGATGTCATTAAGATGAAGAAAGCAATTCATGATCTTGTG 1200
568 GGATGCTGATGATTTTGTGATGTGACCAAAAGATGAAGAAAGCAATGATGATCTGTGGA 509
1201 ACTGTTTGTAAAGAAAACAAGGTTATAGCAGATGTCATATCTTGGGATGTGAAG 1260
508 ACTCATTTGTAGAAAGCAAGGTTGTGAGATGTGACCAAAAGATGAAGAAAGCAATGATG 449
1261 CATTTTCAAGATTTTACGAAAGAGTTGACACCGTTAC---TCACTACTCTTGTGTT 1316
448 CCTTTCAAGCTGATGATTCAGATGTTTGGCCAAAGCAGACGACCTGTTGAGAGTTGT 389

REFERENCE
AUTHORS
Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Corryell, V., Rhana
A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,
Wille, T., Underwood, K., Steptoe, B., Theising, B., Allen, N., Bowers
Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schirk
R., Rutter, E., Kohn, S., Shint, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R., and Wilson, R.
Public Soybean EST project
Unpublished (1999)
JOURNAL
COMMENT
Contact: Shoemaker R./Public Soybean EST Project

FEATURES	High quality sequence stop: 440.
source	location/qualifiers
	1..569

/organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="SOYBEAN CLONE ID: Gm-cl087-5250"
 /clone.lib="Gm-cl087"
 /tissue-type="Roots"
 /lab_host="DH10B"
 /note="Vector: pBluescript II SK+; Site.1: EcoRI; Site.2: XhoI; Site.3: BamHI. cDNA was prepared using polyactric mRNA system kit (Clontech). The cDNA was prepared using the STRATAGENE kit (Clontech). cDNA was synthesized from mRNA using a primer consisting of a poly(dt) sequence with a XhoI restriction site
 5'GAGAGAGAGAGAGAGAGAGACTATGCTGACGTTTTTTTTTTTTTTT)
 EcoRI adapters (5'-OH-AATTGGCAGCAGG and 3'-GCCGTCGCG) were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction sites of the pBluescript vector. The vector was previously dephosphorylated. The ligated cDNA fragments were transformed into DH10B by electroporation. Library is in LB medium with ampicillin and glycerol 8%. Average insert size: 800 bp."

BASE COUNT
 ORIGIN
 172 a 106 c 122 g 169 t

Query Match	11.8%;	Score 203;	DB 14;	Length 569;	
Best Local Similarity	69.6%;	Pred. No. 1,1e-34;			
Matches 275;	Conservative	0;	Mismatches 120;	Indels 0;	Gaps 0;
QY	964	CTACTAAGACAAGAAAGTTATCTCTGAGGATCAGAGGCTAGAACCCCTACTTCTTC	1022		
Db	50	CGTGTAAACAAAGAAAGCTAAGCATGTGATGCATCAGACTCMAAAMACCGAGTCTTGTC	109		
QY	1024	AGAAAGCCCAATTCTATCTATCTCTACAGAGTCCAGCCCAATGGCGCTTGAGCAAGTAATGT	1083		
Db	110	ACAAAGAGACTTTTCTTCTTCTTCTACACACAGAGTCCAGCCCTAATGGCACTAGAAACAAAGTTAT	169		
QY	1084	CTGACCCGGGATTAAGCAGAGATGAAGTCGATGACGATGTGGAGATTTTGAAGATGCCACGA	1143		
Db	170	CAGAGCCGGATATGTGTAAGACCAACATTTATGACGCAATGTGCGATCTTTGAAGATTAAGAGA	229		
QY	1144	TGCTTGATGACTTTTGATGTGATGTGAATTAAGATGAAGCAATTCATGCATCTTTGGAACT	1203		
Db	230	TGCTTGATGATTTTGTGATGTTCCTCCAAAGATGAAGAAAGCACTCAATCTCTGTGAACCT	289		
QY	1204	CGTTTGTAGAAGAAACAAAGGGTTTATAGACAGATGCTATATCTCTTGCGCATGTGAAGCAT	1263		

Db 290 CTTTATATGGAGCAAAAGGGTGGTCGCACATGGTATATTCGGGGCGCTGGAGCAT 349

Qy 1264 TTTTCAGATTTTACAGAAAGAGTGGACCGGTACTACACACTCTTGGGTGGAGAT 1323

Db 350 TTTCCAAAGCTTCATCGAAAAGAGCTGATCTCATCTCTTATTTTGGTACTCTTC 409

Qy 1324 TGTTTTGTATTAACATGGAACCATGGAAGCTTGC 1358

Db 410 TTCCATGTGATTTATCTAATTTTATGAGATTTTC 444

RESULT 9	BI479743	600 bp	mRNA	linear	EST 28-AUG-2001
LOCUS	BI479743				
DEFINITION	BI479743	WHE3451-E10-11925	Wheat pre-methylation spike cDNA library	Triticum aestivum	
ACCESSION	BI479743	WHE3451-E10-119	mRNA sequence.		
VERSION	BI479743.1	GI:15323794			
KEYWORDS	EST.				
SOURCE	bread wheat.				
ORGANISM	Triticum aestivum				

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 600)	Anderson, O.D., Chao, S., Choi, D.W., Close, T.J., Fenton, R.D., Han, P.S., Hsia, C.C., Kang, Y.Y., Lazo, G.R., Miller, R., Rausch, C.J., Seaton, C.L. and Tong, J.C.	The structure and function of the wheat genomes - pre-anthesis spike cDNA library	Unpublished (2000)	Contact: Olin Anderson

US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
10700 B. F. Storer Highway, Albany, CA 94710, USA
Tel.: 5105595773
Fax: 5105595818
Email: candersn@pw.usda.gov
Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20
Seq primer: Stratagene SK primer.
Location/Qualifiers
1. 600

/organism="Triticum aestivum"
 /cultivar="Chinese Spring"
 /db_xref="taxon:4565"
 /clone="WB3451_E10_119"
 /clone_1fb="Wheat pre-anthesis spike cDNA library"
 /clone_type="Spike before anthesis"
 /tissue="Spike before anthesis"
 /dev_stage="Adult plant"
 /lab_host="E. coli SOLR"
 /note="Vector: Lambda Uni-ZAP XR, excised phagemid;
 Site_1: EORI; Site_2: Choi; Plants were grown in the
 greenhouse. Whole spike with awns trimmed, white, green
 and yellow anther were collected and total RNA, and
 poly(A) RNA were prepared, a cDNA library was made, and
 the cDNA clones were in vivo excised to give phagescript
 phagemids in the T7 Close lab (Choi, Close, Fenton) at
 the University of California, Riverside. Plasmid DNA
 preparations and DNA sequencing were performed in the OD
 Anderson lab (all other authors)."

Query Match	11.7%	Score 201.4;	DB 13;	Length 600;
Best Local Similarity	70.0%	Pred. No. 2,5e-44;		
Matches 211; Conservative	0;	Mismatches 116;	Indels 0;	Gaps 0;
OY	1004 TGAAGCCACCTACTTCTTCAGAAGCCCATTCATCTTCACAGAGTCGACCAT	1063		
Db	214 TGCGACGCCAACCTTTGTGCAGAAGCTCAGTTCTTCCTTCACAGGGCAGGCCAT	273		

DEFINITION	AV817525 NC_017410.1	CDNA	CDNA	NC_017410.1	24 NOV 2011	1	
ACCESSION	AV817525	NC_017410.1	CDNA	CDNA	NC_017410.1	24 NOV 2011	1
	mRNA sequence.						

VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

AV817525.1 GI:19859A15
 EST.
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 391)
 Seki, M., Narusaka, M., Ishida, J., Kamuya, A., Setou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinozawa, A., Muramatsu, M., Hayashizaki, Y.,
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 unpublished (2002)
 contact: Motoaki Seki

3-1-1, Koyogadai, Sukunaba, Ibaraki 305-0074, Japan
Tel.: 81-298-36-4359
Fax: 81-298-36-4359
E-mail: yaseb@iact.ibaraki.ac.jp
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). The cDNA library was ligated to modified lambda FIC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our Web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES	
Source	Location/Qualifiers
	1. .391
	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"
	/clone="RAF109-94-M06"
	/clone_11b="RAF19"
	/dev_stage="plants at various developmental stages from germination to mature seeds"
	/lab_host="DH10B"
	/note="Site_1: BamHI; Site_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"
BASE COUNT	102 a 100 c 73 g 116 t
ORIGIN	

Query Match	11.5%;	Score 197.8;	DB 10;	Length 391;
Best Local Similarity	71.7%;	Pred. No. 1.5e-33;		
Matches 259; Conservative	0;	Mismatches 102;	Indels 0;	Gaps 0;

QY	1051	GAGTCACGACATGGCGCTTGAGCAAGTATGCTCAGCGGATGACGATGAAAGTC	1110
Db	388	GAGTCACGCCCTGGCTCTGAGAACAACTCTTCGAGCCGGATGATGAAATGAAGTC	329
QY	1111	ATGAGCATGGTCAGATTTGAGATATGCCAGATGGTATGATCACTTGGATGAGAA	1170
Db	328	ATGATGATGGTCAGATTTGAGATGAAAGATGCTCATGATTTCTGTATGATGACTA	269
QY	1171	AAGATGAAGAACATATCATGATCTGTGGAACATCGTTTGAAGAAACAAAGGTTATAG	1230
Db	268	AAGATGAGAAACATGATGATGATGCATGTGGAGACTCGTTTGTGAGAGAACGCCAGATTAAG	209
QY	1231	CAGATGGGCATATCTCTTGGGCATGTGAGACATTTTCAAGATTTTACGAGAAAGAGTTC	1290
Db	208	CAGATGGTCATCTTCATATGGCATGACGAGGCATTTCCAAAGATTTGACCGGACCCATCATGG	149
QY	1291	ACCGTTACTCATCACTCTCTCGTGTGGAGATGTTTTGATTTAACTATGAGAACATG	1350
Db	148	TTTCGAAACACCGCACTGATTTGTTGGCTGGGAGATGTTTATGTTGAAACTGTGGAAACCAAG	89
QY	1351	GACTGTGACTACGCCACCATCAACACTGCATATCCATCTCGAGATTTGCCGTATA	1410
Db	88	GTCCTTTGGAAGCCCGAAGCATGAAACAACGTATATACCTTTCTCGAACAGCTCCCAATTT	29

Db	28	G	28
RESULT 12			
LOCUS	B0740672		
DEFINITION	B0740672	561 bp	MRNA
	sag50e05.v1	Gm-c1076	Glycine max cDNA clone SOYBEAN CLONE ID:
	Gm-c1076-3374	5'	similar to TR:Q9ZNT9 Q9ZNT9
	FERTILIZATION-INDEPENDENT SEED 2	PROTEIN.	;', mRNA sequence.
ACCESSION	B0740672		
VERSION	B0740672.1	GI:21887459	
KEYWORDS	EST.		
SOURCE	soybean.		
ORGANISM	Glycine max		

REFERENCE
AUTHORS
1 (bases 1 to 561)
Shoemaker, R., Kelm, P., Vodka, L., Expediting, J., Coryell, V., Khana
A., Bolla, B., Marini, M., Hillier, L., Kucaba, T., Matlin, J., Beck, C.,
Wille, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers
Y., Pearson, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurr
R., Ritter, E., Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann
R., Waterston, R. and Wilson, R.
Public Soybean EST project
Unpublished (1999)
Contract: Shoemaker R/Public Soybean EST Project
JOURNAL

FEATURES

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available through: Resgen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: cu@resgen.com web site:
www.resgen.com
Seq primer: -40R from Gldpc
High quality sequence stop: 426.
Location/Qualifiers

```

FEATURES
source
1. Location/Qualifiers
   1.561
   /organism="Glycine max"
   /db_xref="taxon:3847"
   /clone="SOTBEAN CLONE" ID: Gm-cl076-3274"
   /clone_11b="Gm-cl076"
   /tissue_type="wounded cotyledons"
   /dev_stage="11 day old seedlings"
   /lab_host="DH10B"
   /note="Vector: pluescript II SK(+), Site_1: EcoRI, Site_2:
XhoI; The cDNA library was constructed from mRNA isolated
from 11 day old seedlings treated with that were treated
with 2 ugs/ml of a crude glucan elicitor preparation
isolated from the mycelial walls of Phytophthora sojae.
The library was prepared using the Stratagene pluescript
II SK(+) library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a
poly(dT) sequence with an XhoI restriction site. EcoRI
adaptors were ligated to the blunt-ended cDNA fragments
followed by XhoI digestion. The cDNA fragments were
directionally cloned into the EcoRI-XhoI restriction site
of the pluescript vector. The ligated cDNA fragments were
transformed into E. coli Electromax DH10B host cells. Plant
material was provided by Michael G. Hahn (Complex
Carbohydrate Research Center, University of Georgia) and
the library was constructed by Anu Khanna (Ila Vodka lab
, University of Illinois)."
BASE COUNT
166 a
ORIGIN
91 c 123 g 181 t
Query Match      11.0%; Score 190; DB 14; Length 561;
Best Local Similarity 72.2%; Pfd. No. 8; 6e-32;
Matches 247; Conservative 0; Mismatches 95; Indels 0; Gaps 0;

```


FORWARD: AAGGGGATGCTGCAAGCGG
 BACKWARD: GCTCCGGCTGATGTTGTCG
 Seq primer: GGTGTAAACGACGCGCCG
 High quality sequence stop: 364.

FEATURES

source

1..364

/organism="Populus tremula x Populus tremuloides"

/db_xref="taxon:47664"

/clone_lib="Hybrid aspen plasmid library"

/tissue_type="Cambial region"

/dev_stage="1.5 m actively growing tree"

/lab_host="E. coli"

/note="Vector: pBluescript SK; Site_1: SalI; Site_2: NotI; Cambial region tissues, including developing xylem, the meristematic cambial zone and the developing and mature phloem, was harvested from 1.5 m actively growing trees. cDNA was prepared and cloned into lambda gt22a. DNA was isolated and subcloned into pBluescript SK using SalI and NotI restriction enzymes."

BASE COUNT

111 a 68 c 94 g 90 t 1 others

ORIGIN

Query Match 10.2%; Score 176.4; DB 9; Length 364;
 Best Local Similarity 75.5%; Pred. No. 8.5e-29;
 Matches 244; Conservative 0; Mismatches 77; Indels 2; Gaps 2;

951 GCTGTGCTCCCTGCTCTAAGACAGAAAGTTATCTGAGCGCATCAGAGCTAGAAC 1010
 44 GCCATGCTACATATTTCCAAAGACAGAAATTTATCAATTTGAACGCGCTGAGAAC 103
 1011 CACCTACTTCTTCAAGAACGCCAATTTATCTATCTCAGAGCTCAGCAATGCGCTT 1070
 104 CTTACACTCTCTTCAAGAACGCCAATTTATCTATCTCAGAGCTCAGCAATGCGCTT 163
 1071 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 1130
 164 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 223
 1131 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 1190
 224 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 282
 1191 CATCTTGAACCTGTTTATAGAAAGGTTATAGAGATGATGATGATTTT 1250
 283 CACTTGTGACATCATTTTGTGAGGAAGCGGCTTTCAGATGACATATC-CCGGG 341
 1251 GCATGTGAAGCATTTTCAAGATT 1273
 342 GCATGTGAAGCATTTTCAAGATT 364

RESULT 15

BG314044

LOCUS

DEFINITION

405 bp mRNA linear EST 16-Apr-2001

WHE2467.D04.G07Zs Trilicium monoccum early reproductive apex cDNA

sequence.

BG314044

GI:13115847

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Trilicium monoccum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

Triliceae; Trilicium.

1 (bases 1 to 405)

REFERENCE

AUTHORS

Anderson, O.D., Chao, S., Dubcovsky, J., Echenique, V., Han, P.S., Hsiao, C.C., Kang, Y., Lazo, G.R., Miller, R., Kausch, C.J., Seaton, C.L., Stamova, B. and Tong, J.C.

The structure and function of the expressed portion of the wheat genomes - Early reproductive apex cDNA library from Trilicium monoccum unpublished (2001)

COMMENT

Contact: Olin Anderson
 US Department of Agriculture, Agriculture Research Service, Pacific West Area, Western Regional Research Center
 800 Buchanan Street, Albany, CA 94710, USA
 Tel: 5105595773
 Fax: 5105595818
 Email: andersen@wars.ars.gov

Sequence have been trimmed to remove vector sequence and low quality sequence with phred score less than 20

Seq primer: Stratagene SK primer.

Location/Qualifiers

1..405

/organism="Trilicium monoccum"

/cultivar="DV92"

/db_xref="taxon:4568"

/clone_lib="WHE2467.D04.G07"

/tissue_type="Trilicium monoccum early reproductive apex cDNA library"

/tissue_type="Early reproductive apex"

/dev_stage="Seven week-old plants"

/lab_host="E. coli XL0LR"

/note="Vector: Lambda Uni-ZAP XR, excised phagemid; Site_1: EcoRI; Site_2: XhoI; The tissue, total RNA, and poly(A) RNA were prepared from apex at double-ridge stage to terminal-spikelet stage during transition from vegetative state to flower state, a cDNA library was made, and the cDNA clones were in vivo excised at the University of California, Davis (V. Echenique, B. Stamova, J. Dubcovsky). Plasmid DNA preparations and DNA sequencing were performed in the OD Anderson lab (all other authors)."

BASE COUNT

105 a 85 c 104 g 111 t

ORIGIN

Query Match 10.1%; Score 174.6; DB 12; Length 405;
 Best Local Similarity 68.8%; Pred. No. 2.2e-28;
 Matches 240; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

1056 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 1115
 4 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 63
 1116 GATGTTGACATTTTGAAGATGCCAGATGCTGATGATGATGATTTT 1175
 64 GACATTTGCCGACTTGAAGATGACGAGATGCTTGTGATTTTGTGATGACGAAGAT 123
 1176 GAAAGCAATTCATCTTTGAACTGTTGTAGAAACAAAGGTTATAGCAGAT 1235
 124 GAGCAAGTATGCTCTGACCGGATGACGAGATGAGTGAAGTCAGATGATGATTTT 183
 1236 GGTCAATATCTTCTGGGATGATGAGATTTTCAAGATTTTACAGAAAGTTCACCGT 1295
 184 GGGCATATATACCTTGGGCTGTGAGGCTTCTCCGGCTTCATGAGCCGAGCTTACAA 243
 1296 TACTATATCTTCTGTTGTTGAGATTTTGTATTAATATATGAGCAATGAGTAT 1355
 244 AACCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 303
 1356 GTGCAAGTATGCTCTGTTGAGATTTTGTATTAATATATGAGCAATGAGTAT 1404
 304 CTGATGATGCGCGGCGATGACATGACATGACATGACATGACATGACATGACATGAC 352

Search completed: June 19, 2003, 22:08:41

Job time : 2572 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 19, 2003, 20:10:48 ; Search time 1286 Seconds

(without alignments)
9309.105 Million cell updates/sec

Title: US-09-890-220-1

Sequence: 1 caagctcttcatttgct.....aatgattctgctataact 1722

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 7816032 seqs, 3476047932 residues

Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PC1_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq2:*
9: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
10: /cgn2_6/ptodata/2/pna/US10_NEW_COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	366.8	21.3	1722	US-10-425-114-11348	Sequence 11348, A
2	366.8	21.3	1838	US-10-424-599-111422	Sequence 111422, A
3	257.4	14.9	1841	US-10-425-114-31883	Sequence 31883, A
4	236.2	13.7	2640	US-10-231-778-6	Sequence 6, Appl1
5	230.2	13.4	1004	US-10-437-963-11609	Sequence 11609, A
6	227.2	13.2	1070	US-10-424-599-14113	Sequence 14113, A
7	192.4	11.2	716	US-10-424-599-111421	Sequence 111421, A
8	185.6	10.8	443	US-09-837-604B-13507	Sequence 13507, A
9	185.6	10.8	443	US-09-837-604A-13507	Sequence 13507, A
10	174	10.1	558	US-10-389-048-21629	Sequence 21629, A
11	158.8	9.2	2700	US-10-437-963-91042	Sequence 91042, A
12	138.8	8.1	554	US-09-531-113-5697	Sequence 5697, A
13	138.8	8.1	554	US-09-531-113-5697	Sequence 5697, A
14	125.6	7.3	322	US-10-424-599-40645	Sequence 40645, A
15	104	6.0	453	US-09-837-604B-6414	Sequence 6414, A
16	104	6.0	453	US-09-837-604A-6414	Sequence 6414, A
17	104	6.0	453	US-10-437-963-97425	Sequence 97425, A
18	101.2	5.9	2439	US-10-425-114-30709	Sequence 30709, A
19	100	5.8	531	US-09-837-604B-940	Sequence 940, A
20	100	5.8	531	US-09-837-604A-940	Sequence 940, A

C	21	91.8	5.3	632	10	US-10-424-599-73456	Sequence 73456, A
C	22	82.8	4.8	250	5	US-09-531-113-40606	Sequence 40606, A
C	23	82.8	4.8	250	6	US-09-531-113-40606	Sequence 40606, A
C	24	71	4.1	274	10	US-10-389-048-16407	Sequence 16407, A
C	25	69	4.0	921	10	US-10-424-599-108137	Sequence 108137, A
C	26	63.8	3.7	201	13	US-60-466-412-340472	Sequence 340472, A
C	27	58.2	3.4	939	10	US-10-437-963-81637	Sequence 81637, A
C	28	57.8	3.4	6458	9	US-10-231-778-7	Sequence 7, Appl1
C	29	57.2	3.3	241	10	US-10-424-599-52847	Sequence 52847, A
C	30	57.2	3.3	492	9	US-10-144-771-38577	Sequence 38577, A
C	31	56.8	3.3	961	9	US-10-144-771-31796	Sequence 31796, A
C	32	54.2	3.1	8524	9	US-10-311-455-1814	Sequence 1814, A
C	33	53.2	3.1	2892	1	PCR-US01-29288A-255	Sequence 255, App
C	34	53.2	3.1	2892	10	US-10-149-310-255	Sequence 255, App
C	35	53.2	3.1	350	10	US-10-424-599-8668	Sequence 9688, App
C	36	52.2	3.0	8662	9	US-10-240-485-190	Sequence 190, App
C	37	52.2	3.0	867	9	US-10-321-856-340	Sequence 340, App
C	38	51.6	3.0	867	9	US-10-321-856-342	Sequence 342, App
C	39	51.6	3.0	1397	9	US-10-321-856-343	Sequence 343, App
C	40	51.6	3.0	1397	9	US-10-321-856-345	Sequence 345, App
C	41	51.4	3.0	958	9	US-10-144-771-24706	Sequence 24706, A
C	42	51.4	3.0	1178	10	US-10-425-114-16009	Sequence 16009, A
C	43	50.8	3.0	1141	6	US-09-806-708B-22	Sequence 22, Appl1
C	44	50.8	3.0	1141	6	US-09-806-708B-22	Sequence 22, Appl1
C	45	50.6	2.9	3673778	9	US-10-312-841-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1
US-10-425-114-11348

Sequence 11348, Application US/10425114

GENERAL INFORMATION:

APPLICANT: Liu, Jiongong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

APPLICANT: Cao, Yongwei

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 73128

SEQ ID NO 11348

LENGTH: 1722

TYPE: DNA

ORGANISM: Glycine max

FEATURE:

OTHER INFORMATION: Clone ID: 701055296_FLI

US-10-425-114-11348

Query Match

Best Local Similarity 59.8%; Pred. No. 1,le-80;

Matches 757; Conservative 0; Mismatches 422; Indels 87; Gaps 5;

228 AGAATGTGAGCAGATGTCGCGGAAATCTCACCGAGAGATGTTCAACTGAT 287

161 ACCAATGTCGCGGAAATCTCTCCGACACCTGCGGAGAGAAATTCACAGTGAT 220

228 GAGATGTCGTCATATGTAATGTAACCTGTCGACTATATACATCTTCACCTGCT 347

221 GAGAGCTTTTATTTATTTGCAAGCTGTGACCTTTACAAATTTCTTACGCGGTGT 280

348 CTAGGCAACCCATCTCTTCACAGATCTTCAACTACAAATTTGAGCAAGCGCAAA 407

281 CTTCAAAATCCCTCTTTCTTTAGGAGATGTTGCGTTATATAAGCAAGCGGTAA 340

408 AGAAGTCAAGATCTACGATGGATGTTCAACTATATAGCATTTGTAATACACATTA 467

341 AGGAGTGTGAGA---GCAAGAAATGTTGATTTCAATTAATAGGACCGCTACACATTC 397

OY	468	CAGAAAACCTAACTTGGGGAGATGTTCTTCCATTCTCTGATGCGATGGGAGC	527
OY	469	CAGAAAACCTAACTTGGGGAGATGTTCTTCCATTCTCTGATGCGATGGGAGC	528
DB	398	CGAAGAGCTAACTGACCGAAGACTTTCTTCCGCTTTGCTTGATGCACTGGGAAC	457
OY	528	TTCAAGGGGCTGCATTTTCATTGGAATTCATCTCACTGATTTATTTGAAATTCAGTTCAG	587
DB	458	TTTAAAGGTTGGGATTTTCATCTTTGTTTCATCATGATCTATTCACTTGAATCTGG	517
OY	588	CTTTTCCAAATATACCAGACAGTTATGTTCTGTAAACCTTAATTCCTTCATATTTTGA	647
DB	518	GTTACTCATGATATACCAGACAGTGAATGTCCTCTGTAAATAATTTGATATTTGATGACAG	577
OY	648	GAGAAAGAGATGATAGCAT---AAATTTGGGCTCTCTCTCTCTGCTGGAACCTGCT	704
DB	578	AAATTTGCTCTATGGATTAATTTCCACATGTCGCAAACTCTTCTTCTGTCACAGCTCGA	637
OY	705	AAAGGGAGACAAAGAGCTGGCAGAAATTAACCCAGACAGACTTAAAGTATGCTTTTACCG	764
DB	638	AAAGCTGTAAGAAAGAGACTCTGTTCAATTTGAAAAAGCACCAATGTAAATTTCTGGAG	697
OY	765	TTGGATTCACCCA-----	777
DB	698	TTGGACTACCCAGAGAGCATACATAATGATTTCTACAAAAAGATGATATTCATTATCC	757
OY	778	-----GTTTAACTAATGGCACAGAAAATGAAATTCACCTACTTAATGA	820
DB	758	TGCAGAGAGAGAAATGTGTCTAGAACATCTGTAATGTGAAAGATTTTCCATGGGAAGA	817
OY	821	T-----GGAACCGCTGTTTAGGATATCCCGAGGACACAGAGCTGTGTGCACATTT	872
DB	818	AATGACGAGAGAAATTTTGGTCTGATCATCTGGCACATGAGCAACCTGGACATGTG	877
OY	873	GAGATGACCGACCAACATTCACCA-----GCCATAGGCCACCTGTTCTGAGAGCT	923
DB	878	GAATCCAGTTTCAACATTCACAGGTGTTTCAATGGCATGCCCACTCTCTGAGAACCTT	937
OY	924	GGAGCTAAATTTATTTGACAGACAGCTGGGCGCCCTGTACTAAGACACAAAGATTA	983
DB	938	GAATGTGTAAATCAATATGTATAAAGTATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTG	997
OY	984	TTGCTGAGGAGTACAGAGCTGAGAACCCACTCTCTTTCGAAAGCCAACTCTATGAT	104
DB	998	AGCATGGATGATCACTACATCAAGAAACGGAATGCTTCTGCAAGAGACATTTTCTTAC	105
OY	1044	TTCTACAGAGATCCAGCATATGGCGCTGTAGCAAGATTAATGTGACCGGATACGAGAT	110
DB	1058	TCACACAGATTCACGCTATGGCCTACTAGAACAAAGTGTATTCAGACCTGATATGAGAC	111
OY	1104	GAAGTGCATAGCATTTGACATTTTGAATATGCCACAGATGCTTATGACTTTTGGAT	116
DB	1118	GAATTTATATACACATTCAGATATCTTGAAGTATACAGATGCTGACGATTTTGGAT	117
OY	1164	GTGCAATTAAGATGAAAAGCAATTCATGCACTTTTGAATCTGTTTGTAAAGAAAACAAAG	122
DB	1178	GTTTCCAAATATAAAAACAGCTCATGATCTGTGGAACCTTTTATAGAGAAACAAAG	123
OY	1224	GTTATATACAGATGTCATATCTCTTGGGCGATGTGAAGCATTTTCAAGATTTTACGAGAAA	128
DB	1238	GTCTGTACAGATGTCATATCTCTGCGGCGCTGTGAGCATTTTCAAGCTTTCATGAGAAA	129
OY	1284	GATTTACACGTTACTCATCTCTCTTGGGTGGAGATGTTTGGTTTAAACATATG	134
DB	1298	GACCTGATCTCATCTCCACCTTATTTTGGGTGGAGATTATTCATGATCAAAACTTGG	135
OY	1344	AACCATGAGACTGTGCACTACACCCACCAACCAACATGCAATACCATCCGAGAAATGC	140
DB	1358	AATCATGAGCTTCTTGTATGCTCTGTACATGAAACACATGTAGATATTTAGATGTTAC	141
OY	1404	CGTAAAT 1409	
DB	1418	AGGAAT 1423	

```

RESULT 2
US-10-424-599-111422
: Sequence 111422, Application US/10424599
: GENERAL INFORMATION:
: APPLICANT: La Rosa Thomas J
: APPLICANT: Kovalic David K
: APPLICANT: Zhou Yihua
: APPLICANT: Cao Yongmei
: TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-2153223JB
: CURRENT APPLICATION NUMBER: US/10/424,599
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 285684
: SEQ ID NO 111422
: LENGTH: 1838
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1
US-10-424-599-111422

```

Query Match	Similarity	21.3%	Score 366.8	BD 10	Length 1838
Best Local	Similarity	59.8%	Pred. No. 1,16-80		
Matches	Conservative	0	Mismatches 422	Indels 87	Gaps 5
Qy	228	AGAAATGTAGGACGAATTTGTCGGCCGAAATCTCACCGGAGGAAAGATTTCAACTGAT	287		
Db	255	AGCATGTGCGCGCAAAATTTCTCCGTACACCATCCGGTCAGAGAAAGAAATTCGACGTAT	314		
Qy	288	GAAATCTCTGTATATATGTAAACCTCTTGACTATATATACATCTTTCACCTGCTCT	347		
Db	315	GAAGCGCTTATGATTTATTTGGCAAGCCTGTATACATTTTCTTACGCGGTGCT	374		
Qy	348	CTAGGCAACCATCGTTTCTTCCAAAGATGCTTGAATCAACAATTTGAGCAAGCGCAAA	407		
Db	375	CTTCAAAAATCTCTTCTTCTTGAAGGATTTGGCTTATTAATAAGAGCAACCCGTAA	434		
Qy	408	AGAAATCAAGATCTACTGGATGGTACTTTTCACTAATAAGATTTGATAATACACTTA	467		
Db	435	AGAGGTTTGA --- GCAGGAATTTGTGATTTTCAATTATATAGGACCGCTACCAATCTT	491		
Qy	468	CAGAAAATCGAAGTTAGGAGAGATGTCTTGTCCATTTTGCATTCGTATGTGTAC	527		
Db	492	CGAAATACGAAAGTACCGAAGACTTTCTTCCGTTTGTCTGTATGCTAGTGTGAAC	551		
Qy	528	TTCAAGGGGCTCAATTTTCATTTGAATTCATCTCATATTATTTGAATTTGATGTTCAAG	587		
Db	552	TTTAAAGGTTTCGATTTTCATCTTTGTTTCATCATCATATTCATTCACCTTTGAGTTCTGG	611		
Qy	588	CTTTTGSAGATATACAGACAGTTAAATGTTTCTGTAAACCTTAATTCCTTCATTTTGG	647		
Db	612	GTTACTGGAAGATTACCAACAGCTGAATGCTCTGTAAATTTGATATTTGAGATAGG	671		
Qy	648	GAAAGAGAGATGATGACGAT --- AAATTTAGCCCTTCTCTCTGTCTGGAACCTGCT	704		
Db	672	AATGTGTCTGATGAGATTAATTCACAAATCGCAAACTCTTCTCTCTGTCTGACACTGCA	731		
Qy	705	AAGCGAGACAAAGAGGTGGCAGAAATTAACACAGAGACTTAAAGTATGCTTTTACCG	764		
Db	732	AAACGTGAGAAAGAGACTGTCTTCAAAATGAAAAAGCACCAATGTAAATTTCTGGAG	791		
Qy	765	TTTGATTCACCA ---	777		
Db	792	TTTGACTACCGAAGGATACATTAATGATTTTACAAAAGATGATATCTTATCC	851		
Qy	778	-----GTTTAACTAATGGCAGCAAAATGAGATCCACTACTTATATGA	820		
Db	852	TGCAAAGAGAAATGTGTCTAGAAACATCTGTATGATGAAAGATTTTCCCTACTGTGAA	911		
Qy	821	T-----GGAACCGTGTTTATGATATCCGAGGCAACAGAGCTGTGTGACATTT	872		

Db 912 AATGAGGAGGAAAAATTTGGTCTCATCTCGGACCAATGACAACTGGAGATGTG 971
 QY 873 GAGATGACCAAGCAATTCACCA-----GCCATAGCCCACTCTTCTCGAGCGT 923
 Db 972 GAATCCAGTTTCAACATTCACAGGTGTTTCAATTGCCATGCCCAATCTTCTGGACCT 1031
 QY 924 GGTGCTAAAGTTATTTGACAGCAAGAGTGTGCTCCCTGCTACTAAGACAAAGATTA 983
 Db 1032 GAATGTAGTAATCAATATGTATAAAGATCCGTGCTGCTCTAAGCAAGAGATTA 1091
 QY 984 TCTGCTGAGCGATCAAGGCTAGAACCCACTACTCTTCTGAGAAAGCCCAATTCATCAT 1043
 Db 1092 AGCATGATCGATCAACACTCAAGAAACCGAATGCTCTGAGAAAGAGACTTCTTTCAC 1151
 QY 1044 TCTCAGAGATCCAGCAATGGCGCTGAGCAAGTATGCTGACCGGGATGACGAGAT 1103
 Db 1152 TCACACAGAGTCCAGCTTATGGCAGCTAGAACAGTGTATCAGACCGTGTACTGAGAC 1211
 QY 1104 GAATGATGAGCATGTTGAGATTTGAAAGATGCGCAGATGCTGATGATCTTGTGAT 1163
 Db 1212 GAAGTGTATGAGCAATTCAGATCTTGAAGATAGAAGATGCTGACGATTTGTGAT 1271
 QY 1164 GTGAATAAAGATGAAAGCAATTCATGCACTTTGCACTCGTTTGAAGAAACAAAG 1223
 Db 1272 GTTTCGAAAGATGAAAGCAAGCTGATGCACTCTTGTGAGAAAGCAAG 1331
 QY 1224 GTTATGAGATGATGATATCTCTGCGATGTGAGCAATTTCAAGATTTTCAGGAA 1283
 Db 1332 GTGCTAGCAGATGATGATATCTCTGCGATGTGAGCAATTTTCAGGAA 1391
 QY 1284 GAGTGGACACGTTACTCATCTCTGCTGTTGGAGATGTTTGTATTAACATGAG 1343
 Db 1392 GAGCTGATCTATCTCAGCTTATTTGGTGTGAGGTTATCATGATCAAACTTGG 1451
 QY 1344 AACCTGACTGTGCTGATCAGCAGCCACCAATCAACTGCAATCACTCTGAGATTCG 1403
 Db 1452 AATCATGATCTCTGATGCTGATGATGATGATGATGATGATGATGATGATGAT 1511
 QY 1404 CGTAAAT 1409
 Db 1512 AGAAAT 1517
 RESULT 3
 US-10-425-114-31883
 ; Sequence 31883, Application US/10425114
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jingdong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaska, Jack E.
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53313)B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 31883
 ; LENGTH: 1841
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: UC-ZMFLB7322612_FLI
 US-10-425-114-31883

Query Match 14.9%; Score 257.4; DB 10; Length 1841;
 Best local Similarity 56.5%; Pred. No. 1.8e-53;
 Matches 573; Conservative 0; Mismatches 391; Indels 51; Gaps 3;
 QY 424 CTGGATGGTACTTTCAACTATAGGATGTATATACACATTACGAAAGTGAAGTTA 483
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

Db 493 CTGAAATGACTTTTTTATTTACAGTACTACAAATACATACATGCAAAAGCTGAAGTTA 552
 QY 484 GGGAGATGTGTTGTCATTTGGCTATGCTATGATGTGTACTGCTTCAAGGGCTGCAT 543
 Db 553 CTGAAATTTTCTGCTCCATTTGCTGTTGCTGATGATGATGATGATGATGATGAT 612
 QY 544 TTTCAATTTGATTCATGATGATTTTGAATTTGATGATGATGATGATGATGATGAT 603
 Db 613 GGCATTAATTTATCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 672
 QY 604 AGCACTTAATTTGATTTGTAATACTTAATTTCTTCAATTTTGAAGAAAGAGATGAT 663
 Db 673 AGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 722
 QY 664 AGCATTAATTTGAGCCCTCTCTCTGCTGCTGCAAGCTGTAGCGGAGCAAGAGATG 723
 Db 733 TTGATCCAAAGCATCAAACTTTTCTCTGCTGCTGAGGTTTGAAGAGCTGAGAG--G 788
 QY 724 GCAGAAATACACAGAGACTTAAAGTATGCTTTTACCGTTGATTCACCGATTTAA 783
 Db 789 TTGGAAGACACAGCTGAGAAATTCAGGCAATTCACCCATATTTATGATCAGAT-- 846
 QY 784 CTATGAGCACAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 843
 Db 847 -----CACATGAATGATGCCAGGAGGCTGAAAGATG 879
 QY 844 CCGAGGACACAGAGCTTGTGACATTTGATGATGATGATGATGATGATGATGATGATGAT 903
 Db 880 ACTATGTCAGAAAGGAAATGGGCTTTCTG-----TAGCAATATCTT 921
 QY 904 CCCACTCTCTCTGAGAGCTGCTGCTAAAGTTATTTGACAAAGCAAGCTGTGCTGCT 963
 Db 922 CAGTTGACCTATGACGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 981
 QY 964 CTATGAGACAGAAATTTATGCTGAGGATGAGAGGCTAGAGGCTGATGATGATGATGAT 1023
 Db 982 TTGGAAGACAAAGAAATCTATGCTGAGGATGATGATGATGATGATGATGATGATGAT 1041
 QY 1024 AGAAAGCCAAATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1083
 Db 1042 AAAAGGTCAGTCTTCCATTCACACAGGAGCAAGCAATGCTATGGAACAGTCTCT 1101
 QY 1084 CTGACCGGATGACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1143
 Db 1102 CAGATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1161
 QY 1144 TGTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1203
 Db 1162 TGTCTACAGATTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
 QY 1204 CGTTGTGAAGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1263
 Db 1222 CATTTGTGAAAGCAAGAGGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1281
 QY 1264 TTTCAAGATTTTACGAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1323
 Db 1282 TCTCCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1341
 QY 1324 TGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1383
 Db 1342 TCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1401
 QY 1384 ATACCATGCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1438
 Db 1402 ACACAAATCTTCAAGGCTACCAAGAGGAGATGCGGAGCCCAAGTAAATCTACAA 1456

RESULT 4
 US-10-231-778-6
 ; Sequence 6, Application US/10231778
 ; GENERAL INFORMATION:
 ; APPLICANT: Billoreau, Pierre
 ; APPLICANT: Chaudhury, Abdul M.

```

APPLICANT: Dennis, Elizabeth S.
APPLICANT: Koltunow, Anna M.G.
APPLICANT: Luo, Ming
APPLICANT: Pascock, William J.
TITLE OF INVENTION: Method for inducing seed development by down-regulating
FILE REFERENCE: 72-98A
FILE REFERENCE: expression of the F1S2 gene
CURRENT APPLICATION NUMBER: US/10/231,778
PRIOR FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: 09/398,237
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/101,184
PRIOR FILING DATE: 1998-09-21
PRIOR APPLICATION NUMBER: AU PP6061
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6062
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6063
PRIOR FILING DATE: 1998-09-22
PRIOR APPLICATION NUMBER: AU PP6063
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: AU P01345
PRIOR FILING DATE: 1999-07-01
PRIOR APPLICATION NUMBER: AU P01346
NUMBER OF SEQ ID NOS: 239
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO: 6
LENGTH: 2640
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
NAME/KEY: misc-feature
LOCATION: (1)-(2439)
OTHER INFORMATION: Nucleotides from 1 to 2439 represent protein
US-10-231-778-6

Query Match      13.7%; Score 236.2; DB 9; Length 2640;
Best Local Similarity 71.0%; Pred. No. 3.7e-48;
Matches 313; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

954 GTGTCCTCTGCTACTAAGCAAGAAAGTATCTGCTGAGCGATCAGAGCTAGAGCCAC 1013
1915 GTGACTGCTGTAGTGAAGAAAGAGTTACATGAGAGCGGTGCGAGCTTAAAGATG 1974
1014 CTACTCTTGAAGAACGCCAATTCATCTCTACAGAGTCCAGCCCAATGCGCTTGAG 1073
1975 GAGCGCTTAAAGGTGACAGAGTCTATCTACCCAAACAAATGACGCCCAATGCTTGGAA 2034
1074 CAAGTAATGCTGACCGGATAGCGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1133
2035 CAAGTAATGCTGACAGAGATAGCGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1193
1134 GATGCCAGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1193
2095 GAAAGCGCTGAGACTTAAAGCTGTGTGGGTGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2154
1194 CTTTGAAGCTGTTTGAAGAAACAAAGAGTATAGAGATGATGATGATGATGATGATGATGATGATGAT 1253
2155 CTTTGAAGCTGTTTGAAGAAACAAAGAGTATAGAGATGATGATGATGATGATGATGATGATGATGAT 2214
1254 TGTGAAGCAATTTCAAGATTTTGAAGAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1313
2215 TGTGAAGAGTTCCTCAAACTTCAATGAAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 2274
1314 TGTGGAGATGTTTGTGATTTGAATGAATGAAGCAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1373
2275 TGTGGAGATGTTTGTGATTTGAATGAATGAAGCAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 2334
1374 AACAACTGCAATTCATCTCTC 1394
2335 CACAAATGCACTACATCTCTC 2355

```

```

RESULT 5
US-10-437-963-11609/C
Sequence 11609, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Barabazuk, Andrey A.
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO: 11609
LENGTH: 1004
TYPE: DNA
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609

Query Match      13.4%; Score 230.2; DB 10; Length 1004;
Best Local Similarity 68.8%; Pred. No. 8.6e-47;
Matches 316; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

950 AGCTGTGCTCCCTGCTACTAAGCAAGAAAGTATCTGCTGAGCGATCAGAGGCTAGAG 1009
782 AACAGTGTAAAGTGTGAGAGAAAGAGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 723
1010 CCACCTACTCTTGAAGAACGCCAATTCATCTCTACAGAGTCCAGCCCAATGCGCT 1069
722 TCGGAGGCTCTTCAAAAACGCCAGATCTTCTCTCTCAGCGCTCAACCAATGGCAAT 663
1070 TGAGCAAGTAAATGCTGACCGGATAGCGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1129
662 GGAGCAAGTAAATGCTGACCGGATAGCGAGATGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 603
1130 TGAAGATGCGCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
602 TGAAGATGCGCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 543
1190 GCATCTTGAAGTCTGTTTGAAGAAACAAAGGTTATAGCAAGATGATGATGATGATGATGATGATGAT 1249
-542 GCATCTTGAAGTCTGTTTGAAGTCTGTTTGAAGTCTGTTTGAAGTCTGTTTGAAGTCTGTTTGAAG 483
1250 GGCATGTAAGCAATTTTCAAGATTTTCAAGAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 1309
482 GGCATGTAAGCAATTTTCAAGATTTTCAAGAAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAG 423
1310 CTGTGTGAGATGTTTGTGATTTGAATGAATGAAGCAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 1369
422 ATGGTGTGAGAGTGTGATTTTGAATGAATGAAGCAATGAGAGTGTGAGTGTGAGTGTGAGTGTGAG 363
1370 CATCAACAACTGCAATTCATCTCTGAGAGATTTCCCTGAA 1408
362 CATGAATGCTGCAACCAATCTTGAAGGCTACCTGAA 324

RESULT 6
US-10-424-599-141143/C
Sequence 141143, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO: 141143
LENGTH: 1004
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-424-599-141143/C

```



```

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 141143
; LENGTH: 1070
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1
US-10-424-599-141143

```

```

Query Match      13.2%; Score 227.2; DB 10; Length 1070;
Best Local Similarity 68.6%; Pred. No. 4,9e+46;
Matches 313; Conservative 0; Mismatches 143; Indels 0; Gaps 0;

```

```

QY 949 AAGCTGTGCTCCCTGCTACTAAGACAAAGTATCTGCTGAGGAGTACAGAGGCTAGAA 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 824 AAGCAGTGTGCTGAGTGGCCAGACAGAGAGTGTGCAATGAGCGCTCCGACCCAGAA 765
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1009 GCCACCTACTCTTGAGAAAGCCAACTATCATCTCCAGAGTCCAGCCAAAGGCCG 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 764 ACAGTACCTCTTGAGAGAGCGACATTTTTCATTCACACAAAGCTCAGCCAAAGGCAA 705
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1069 TTGACCAAGTAAATGCTGACCCGAGTACGAGAGTAAAGTCCAGATGATGTCAGATT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 704 TTGAACAAGTTTATCCGATTAAGATAGGAGAGATGAAGTGTGATGATGTTGCCGATT 645
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1129 TTGAAGATGCCAGATGCTGTGATGCTTTGTGATGTGAATAAAGATGAAGAAACATTTCA 1188
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 644 TTGAAGATCGAAGAGTGTGAAAGTGTGATGTGAGCAATGAGAAAGACTTTCA 585
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1189 TGCATCTTGGAACTGTTTGAAGAAACAAAGGTTTACAGATGCTCATCTCTT 1248
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 584 TGCATATGTGAACTGATTTGTCGAGACATGCTGTGATTCACATGCTCATTTCAT 525
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1249 GGGCATGTGAACATTTTCAAGATTTTACGAGAAAGTGTGACCGTTACTCATCTCT 1308
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 524 GGGCATGTGAAGATTTTCAAGATTTTACGAGAAAGTGTGACCGTTACTCATCTCT 465
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1309 TGTGCTGTGAGATGTTTGTGATTAACATATGGAACCATGAGTGTGACTCAGCCCA 1368
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 464 CAGGCTGTGAGATTTTATGATGTCACAAATTTACATCATGCTCTTACAGATGTCGGA 405
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1369 CCATCAACAACATGCAATACCATCTCGAGAAATTCGC 1404
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 404 CCATCAATGACTGTATATTTATTTTGAACCAATACC 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-10-424-599-111421

```

```

; Sequence 111421, Application US/10424599

```

```

; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111421
; LENGTH: 716
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1
US-10-424-599-111421

```

```

Query Match      11.2%; Score 192.4; DB 10; Length 716;
Best Local Similarity 68.5%; Pred. No. 2e-37;

```

```

Matches 281; Conservative 0; Mismatches 126; Indels 3; Gaps 1;
QY 228 AAGATGTGAGCAGATTTGTCGGGAAATCTCTACCCGAGAGATGATTTCACTGAT 287
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 214 AAGATGTGAGCAGATTTGTCGGGAAATCTCTACCCGAGAGATGATTTCACTGAT 273
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 288 GAGATCTCTGATATATTTGTAACCTGTTGCACTATTAACATCTTACCTGGCTCT 347
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 274 GAGATCTCTGATATATTTGTAACCTGTTGCACTATTAACATCTTACCTGGCTCT 333
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 348 CTAGCAACCCATCTGTTCTTCCAGATGCTTGAACCTACAAATAGGACAAAGCCGCAA 407
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 334 CTCAAAATCTCTCTTCTTCTTGAAGATGTTTGTCTTATTAATAAAGACAGCGCTATA 393
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 408 AGAAGTCAAGATCTCTCTGAGATGATGATTTTCACTATTAAGATTTGTAATACATTA 467
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 394 AGAAGTCAAGATCTCTCTGAGATGATGATTTTCACTATTAAGATTTGTAATACATTA 450
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 468 CAGAAATCTGAAGTATGAGAGATGTTCTTGTCCATTTTGTCTATGCTATGATGATG 527
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 451 CGGAAATCTGAAGTATGAGAGATGTTCTTGTCCATTTTGTCTATGCTATGATGATG 510
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 528 TTCAAGGGCTGCAATTTTCAATTTGATTCATCATGATTTATTTGAATTTGATGATG 587
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 511 TTGAAGGGTTCGATTTTCATCTTGTTCATCATGATTTATTTCAATTTGGTCTCG 570
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 588 CTTTCGAGATATACAGACATGATTTGTTCTGTAACTTAATTCCTT 637
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 571 GTTACTGAAGATTTACCAAGCAGTGAATGCTCCGTAAATAATTAATTAAT 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-09-837-604B-13507

```

```

; Sequence 13507, Application US/09837604B

```

```

; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: De la Pena, Robert C.
; APPLICANT: Bought, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604B
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
US-09-837-604B-13507

```

```

Query Match      10.8%; Score 185.6; DB 5; Length 443;
Best Local Similarity 72.0%; Pred. No. 8.3e-36;
Matches 242; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

```

```

QY 950 AGCTGTGCTCCCTGCTACTAAGACAAAGATTAATCTGCTGAGCGATCAGAGCTAGAG 1009
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 98 AACAGTGTACAGTTTGGGAAGAACAAAGAGCTGTCTGTTGAACGAGCTGATCCAGAAA 157
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1010 CCACCTACTTCTTCAAGAAAGCCAAATTTATATTTCTACAGAGTCCAGCAATGGCGCT 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 TCGGAGCTCTTCAAAAACGCCAGATTTCTTCACTCAGAGGGCTCAACCAATGGCAT 217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1070 TGAGCAAGTATGCTGACCGGAGATGAGAGATGAAGTCAAGTCAAGTGTGAGATTT 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 218 GAGCAAGTATTTCTGAGATGATGATGAAGATGAAGTGAATGATGATGATGATGATTT 277
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1130 TGAAGATGCCAGATGCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

Db 278 TGAAGATAGAGAAAGTCTGATGATTTGTTGATGTTACAAAAGCAGAAACTTATAT 337
OY 1190 GCATCTTTGGAACCTGTTTGTAGAAAAGAGGTTATAGCAGATGTCATATCTTGT 1249
Db 338 GCATATGTGGAATTCATTTGTTGCGAAACAAAGGCTACTAGCGATGCGCATATTCCTG 397
OY 1250 GGCATGTGAAGCATTTTTCAGATTTTACGAGAAAGA 1285
Db 398 GGCATGCGAAGCATTTTCTCGAGTTTCATGACAAAGA 433

RESULT 9.

US-09-837-604A-13507
Sequence 13507 Application US/09837604A
GENERAL INFORMATION:
APPLICANT: Byrum, Joseph R.
APPLICANT: La Rosa, Thomas J.
APPLICANT: Shukla, Hridayadhiranjan
APPLICANT: De La Pena, Robert C.
APPLICANT: Bougri, Olegs
TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
FILE REFERENCE: 38-21(51892)B
CURRENT APPLICATION NUMBER: US/09/837,604A
CURRENT FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: US 60/197,872
PRIOR FILING DATE: 2000-04-19
NUMBER OF SEQ ID NOS: 81288
SEQ ID NO 13507
LENGTH: 443
TYPE: DNA
ORGANISM: Oryza sativa nipponbare
FEATURE:
OTHER INFORMATION: Clone ID: LIB3432-029-p1-K1-H2
US-09-837-604A-13507

Query Match 10.8%; Score 185.6; DB 7; Length 443;
Best Local Similarity 72.08; Pred. No. 8.3e-36;
Matches 242; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

OY 950 AGCTGTGGTCCCTGCTACTAGAACAAAGATTATGCTGAGCGATCAGAGGCTAGAG 1009
Db 98 AACAGTGTACAGTTGGGGAAGCAAGAAAGCTGCTGTTGAACAGGCTGATCCAGAAA 157
OY 1010 CCACCTACTCTTTCAGAAAGCCCAATTTCTATCTCTCACAGAGTCCAGCCCAATGGCGCT 1069
Db 158 TCGGCACTCTCTCAAAAAGCCAGCTTTTCTCATCTCTCAGGGCTCAACCAATGGCAT 217
OY 1070 TGAACAGTATGTCGACCGGATAGCAGATGAGATGATGATGATGATGATGATGAT 1129
Db 218 GAGCAGAGTTTCTCAGATGATGATGATGATGATGATGATGATGATGATGATGAT 277
OY 1130 TGAAGATGCGCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1189
Db 278 TGAAGATAGAAAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 337
OY 1190 GCATCTTTGGAACCTGTTTGTAGAAAAGAGGTTATAGCAGATGTCATATCTTGT 1249
Db 338 GCATATGTGGAATTCATTTGTTGCGAAACAAAGGCTACTAGCGATGCGCATATTCCTG 397
OY 1250 GGCATGTGAAGCATTTTTCAGATTTTACGAGAAAGA 1285
Db 398 GGCATGCGAAGCATTTTCTCGAGTTTCATGACAAAGA 433

RESULT 10

US-10-389-048-21629
Sequence 21629 Application US/10389048
GENERAL INFORMATION:
APPLICANT: Havukkala, Ilka
APPLICANT: Shenk, Michael A.
TITLE OF INVENTION: Polynucleotides, Materials Incorporating

FILE OF INVENTION: Them, and Methods for Using Them
FILE REFERENCE: 11000.1041U1
CURRENT APPLICATION NUMBER: US/10/389,048
CURRENT FILING DATE: 2003-03-13
NUMBER OF SEQ ID NOS: 25129
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 21629
LENGTH: 558
TYPE: DNA
ORGANISM: Eucalyptus grandis
US-10-389-048-21629

Query Match 10.1%; Score 174; DB 10; Length 558;
Best Local Similarity 62.8%; Pred. No. 6.8e-33;
Matches 304; Conservative 0; Mismatches 175; Indels 5; Gaps 2;

OY 221 TTGAACAAAGATGTGAGCAGAAATGTCGCGGAAATCCTCACCGAGGAAAGTATTC 280
Db 64 TTGGGTTCAATGTGCTCCATCAAGTTCTTGTGAACATTTTGTGATGAAGATGCAACTGC 123
OY 281 AACTGATGAGATCTCTGATATATGTTAAACCTGTTGACATATATACATCTTCACT 340
Db 124 AGCAGAGAGAAATCTCGGTGATTCAGACCTGTTGAGCTATACACATCTTCACTG 183
OY 341 TCGCTCTGAGCAACCATGCTTCTTCCAGATGCTGTAACATAAATGAGCAAA 400
Db 184 CCGTCCATACATATATCTTCTTCTCCGAGATGTTGCGTTCAAAATACAAAGCAAG 243
OY 401 GCGCAAAAGAAAGTCAAGATCTA--CTGGATGATGATTTTCACTATAGATTTAA 457
Db 244 ACGTAAAGGTAATTAAGGTTGAGCGCTGCAATATGATATATATAGGACCGCA 303
OY 458 TAAACATTAAGAAAGTGAAGAGTGAAGAGTGTGCTGCTGCTGCTGCTGCTGCT 517
Db 304 CAAGTGTGATCAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 363
OY 518 ATGTGATGCTTCAAGGGGCTGCAATTTCAATTTGATTTCAATTTGATTTGATTT 577
Db 364 ATGTCAAGCTTTTAAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAGGCTTCAAG 423
OY 424 CGAGTTTGGTATTTGAAGATGATGAGGATGATGATGATGATGATGATGATGATGAT 483
Db 638 CATATTTGAG--AAGAAGAGATGATGATGATGATGATGATGATGATGATGATGAT 695
OY 484 GAGATCCGAGGTTCTGACTGAAAGCAACCTAAATGAGCATTTCTTCTGCGCA 543
OY 696 AAC 699
Db 544 AAC 547

RESULT 11

US-10-437-963-91042
Sequence 91042 Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David R.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 91042
LENGTH: 2700

;; CURRENT FILING DATE: 2003-04-28
;; NUMBER OF SEQ ID NOS: 285684
;; SEQ ID NO 40645
;; LENGTH: 322
;; TYPE: DNA
;; ORGANISM: Glycine max
;; FEATURE:
;; NAME/KEY: unsure
;; LOCATION: (1)..(322)
;; OTHER INFORMATION: unsure at all n locations
;; FEATURE:
;; OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1
US-10-424-599-40645

Query Match 7.3%; Score 125.6; DB 10; Length 322;
Best Local Similarity 61.2%; Pred. No. 6.3e-21;
Matches 197; Conservative 0; Mismatches 125; Indels 0; Gaps 0;

OY 1062 ATGGCGCTTGACCAAGTAATGTCTGACCGGATAGCGAGATGAAGTCGATGACGATGTT 1121
DB 322 ATGGCANTAGAAATTAAGTCTACCCACCGCAGACAGTCAGCAAGATTGACACGACCG 263
OY 1122 GCAGATTTTGAAGATCGCCAGATGCTTATGATGATGATGATGATGATGATGATGAT 1181
DB 262 GACGATGTTGAAGATCGCCAGATGCTTATGATGATGATGATGATGATGATGATGATGAT 203
OY 1182 CAATTCATCATCTTTGGAACCTGTTGTAGAAAACAAAGGTTATAGCAGATGTCAT 1241
DB 202 TAGTCATCATCTTTGGAACCTGTTGTAGAAAACAAAGGTTATAGCAGATGTCAT 143
OY 1242 ATCTCTTGCGCATGTAAGCATTTTCAAGATTTTACGAAAGAGTTGACACCGTTACTCA 1301
DB 142 GTTGTGTTGGCTTGTGAGCATTTTCCAGTTTATGAGAAAAGAGCTGATCTATCTCT 83
OY 1302 TCACCTCTGCTGTTGAGATGTTTGTATTAACATGAGAACCATGACCTTGTGAC 1361
DB 82 GCTTATTTTGGCGTTGAGGATTTTGTATGATCAAACTTGTGAAATTTGCTCTTCTCAC 23
OY 1362 TCAGCCACCATCAACCACTGCA 1383
DB 22 GCGCGTACTTATGATTAATGTA 1

RESULT 15
US-09-837-604B-6414
;; Sequence 6414, Application US/09837604B
;; GENERAL INFORMATION:
;; APPLICANT: BYrum, Joseph R.
;; APPLICANT: La Rosa, Thomas J.
;; APPLICANT: Shukla, Hridayabhiranjan
;; APPLICANT: De La Pena, Robert C.
;; APPLICANT: Bougri, Oleg
;; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
;; FILE REFERENCE: 38-21(51892)B
;; CURRENT APPLICATION NUMBER: US/09/837,604B
;; CURRENT FILING DATE: 2001-04-18
;; PRIOR APPLICATION NUMBER: US 60/197,872
;; PRIOR FILING DATE: 2000-04-19
;; NUMBER OF SEQ ID NOS: 81288
;; SEQ ID NO 6414
;; LENGTH: 453
;; TYPE: DNA
;; ORGANISM: Oryza sativa nipponbare
;; FEATURE:
;; OTHER INFORMATION: Clone ID: LIB3431-030-PI-K2-G11
US-09-837-604B-6414

Query Match 6.0%; Score 104; DB 5; Length 453;
Best Local Similarity 66.5%; Pred. No. 1.6e-15;
Matches 149; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

OY 424 CTGGATGATGATTTCAACTATTAAGATTTGTAATACACATTACAGAAAACGAAAGTTA 483

DB 173 CTGGCAATGCTCTTTTAACTACAAATAATTAATTAATTAATTAATTAATTAATTAATTAAT 232
OY 484 GGGAGGATTTGTTGTCATTTTGTCTATGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 543
DB 233 CTGAAGATTTTCTGCGCAATTTTGTGCTATGCTATGCTATGCTATGCTATGCTATGCTAT 292
OY 544 TTCAATTTGAATTTCAATGATTTATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 603
DB 293 GTCACTTAACGATTCGATGACCTTTTCCATTATAGTTTGGATATCTGAAAGTGCC 352
OY 604 AGACAGTTAATGTTTCTGTAAACTTAATTTCTTCAATTTGAG 647
DB 353 AGGCTGTAAATGTTATGCTGAAAGCTGATTTCTTGGAAACGAG 396

Search completed: June 19, 2003, 23:39:28
Job time : 1299 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:22:23 ; Search time 72 Seconds

(without alignments)
823.563 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRONCRAKSSPEVISTDE.....INNNVNDKDNHSDKYIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A.Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2378	100.0	445	21	VRN2 polypeptide.
2	2285	96.1	440	21	VRN2 polypeptide.
3	1923	80.9	498	21	Arabidopsis thaliana
4	1796	75.5	367	21	Arabidopsis thaliana
5	1635	68.8	339	21	Arabidopsis thaliana
6	923	38.8	611	21	Arabidopsis thaliana
7	906.5	38.1	604	21	Arabidopsis thaliana
8	549	23.1	107	21	Arabidopsis thaliana
9	510.5	21.5	813	21	Arabidopsis thaliana
10	451.5	19.0	186	21	Arabidopsis thaliana

11	368.5	15.5	108	21	Arabidopsis thaliana	Arabidopsis thaliana
12	327	13.8	63	21	Arabidopsis thaliana	Arabidopsis thaliana
13	324	13.6	145	21	Arabidopsis thaliana	Arabidopsis thaliana
14	324	13.6	174	21	Arabidopsis thaliana	Arabidopsis thaliana
15	311	13.1	154	21	Arabidopsis thaliana	Arabidopsis thaliana
16	297	12.5	67	21	Arabidopsis thaliana	Arabidopsis thaliana
17	225	9.5	388	22	Arabidopsis thaliana	Arabidopsis thaliana
18	221	9.3	803	21	Arabidopsis thaliana	Arabidopsis thaliana
19	219	9.2	955	22	Arabidopsis thaliana	Arabidopsis thaliana
20	153	6.4	289	22	Arabidopsis thaliana	Arabidopsis thaliana
21	131.5	5.5	1245	21	Arabidopsis thaliana	Arabidopsis thaliana
22	130.5	5.5	710	22	Arabidopsis thaliana	Arabidopsis thaliana
23	129	5.4	22	21	Arabidopsis thaliana	Arabidopsis thaliana
24	128.5	5.4	1817	21	Arabidopsis thaliana	Arabidopsis thaliana
25	127.5	5.4	700	21	Arabidopsis thaliana	Arabidopsis thaliana
26	127.5	5.4	847	21	Arabidopsis thaliana	Arabidopsis thaliana
27	127.5	5.4	953	21	Arabidopsis thaliana	Arabidopsis thaliana
28	127.5	5.4	953	21	Arabidopsis thaliana	Arabidopsis thaliana
29	127.5	5.4	961	21	Arabidopsis thaliana	Arabidopsis thaliana
30	125.5	5.3	1501	23	Arabidopsis thaliana	Arabidopsis thaliana
31	124.5	5.2	686	21	Arabidopsis thaliana	Arabidopsis thaliana
32	123.5	5.2	2010	21	Arabidopsis thaliana	Arabidopsis thaliana
33	122	5.1	190	14	Arabidopsis thaliana	Arabidopsis thaliana
34	122	5.1	190	14	Arabidopsis thaliana	Arabidopsis thaliana
35	122	5.1	256	13	Arabidopsis thaliana	Arabidopsis thaliana
36	122	5.1	1714	21	Arabidopsis thaliana	Arabidopsis thaliana
37	121	5.1	190	14	Arabidopsis thaliana	Arabidopsis thaliana
38	121	5.1	590	12	Arabidopsis thaliana	Arabidopsis thaliana
39	120.5	5.1	1436	21	Arabidopsis thaliana	Arabidopsis thaliana
40	120	5.0	190	14	Arabidopsis thaliana	Arabidopsis thaliana
41	120	5.0	2380	21	Arabidopsis thaliana	Arabidopsis thaliana
42	119.5	5.0	190	14	Arabidopsis thaliana	Arabidopsis thaliana
43	119	5.0	190	14	Arabidopsis thaliana	Arabidopsis thaliana
44	119	5.0	190	14	Arabidopsis thaliana	Arabidopsis thaliana
45	119	5.0	190	14	Arabidopsis thaliana	Arabidopsis thaliana

ALIGNMENTS

RESULT 1	AA000060	standard; Protein; 445 AA.
ID	AA000060	
AC	AA000060	
XX	16-NOV-2000	(first entry)
DE	VRN2 polypeptide.	
XX	VRN2 polypeptide.	
XX	VRN2 polypeptide.	
KW	leaf size; leaf shape; shade avoidance response; reproduction;	
KW	breeding; pollination; cultivation.	
XX	Arabidopsis thaliana var 'Landsberg erecta'.	
OS	Arabidopsis thaliana var 'Landsberg erecta'.	
XX	Arabidopsis thaliana var 'Landsberg erecta'.	
FT	Key	Location/Qualifiers
FT	Modified-site	57..61
FT	Region	/note="Putative NLS"
FT	Modified-site	90..111
FT	Modified-site	/label="Zinc finger motif"
FT	Modified-site	158..159
FT	Modified-site	/note="Putative NLS"
FT	Modified-site	170..174
FT	Domain	/note="Putative NLS"
FT	Domain	281..316
FT	Misc-difference	323
FT	Misc-difference	/note="Putative acidic activation domain"
FT	Misc-difference	323
FT	Misc-difference	/note="mutation in mutant gene vin2-1 causes a change from a Trp codon to a stop codon resulting in a truncated mutant protein"
PN	WO200044918-A1.	

```

XX 03-AUG-2000.
PD
XX
XX 28-JAN-2000; 2000WO-GB00248.
XX
XX 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Dean C, Gendall A;
XX WPI: 2000-499333/44.
XX N-PSDB; AAA47751.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
XX with altered vernalization response, flowering time, leaf size and/or
XX shape or shade avoidance response for maximized reproductive success
XX
XX Claim 49; Fig 6; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX plant encode polypeptides which are capable of affecting one or more
XX vernalization responses such as, flowering time, leaf size and/or
XX shape or the shade avoidance response of a plant into which the
XX nucleic acid is introduced. Introducing such sequences into plants
XX to alter these characteristics maximises the reproductive success of
XX the plant.
XX
XX Sequence 445 AA;
SQ
Query Match 100.0%; Score 2378; DB 21; Length 445;
Best Local Similarity 100.0%; Pred. No. 2,4e-234;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSFPLPCINRYKIGAKRR 60
DB 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSFPLPCINRYKIGAKRR 60
QY 61 KSRSTGMVYVFNKDCNNITLQKTEVREDSCSPFCSMGSGFKGLQFHLNSHDLFEFEFKL 120
DB 61 KSRSTGMVYVFNKDCNNITLQKTEVREDSCSPFCSMGSGFKGLQFHLNSHDLFEFEFKL 120
QY 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHSSLDAGAKYILTSE 240
QY 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHSSLDAGAKYILTSE 240
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
QY 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
QY 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
DB 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
QY 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
DB 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420
QY 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420
DB 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420
QY 421 DHPNDJNNKNNVNNKNNNSRDKYIK 445
DB 421 DHPNDJNNKNNVNNKNNNSRDKYIK 445

```

```

DT 16-NOV-2000 (first entry)
XX
XX VRN2 polypeptide.
DE
XX
XX Vernalization gene: VRN2; plant characteristic; flowering time;
XX leaf size; leaf shape; shade avoidance response; reproduction;
XX breeding; pollination; cultivation.
XX Arabidopsis thaliana (Columbia).
XX WO200044918-A1.
XX
XX 03-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-GB00248.
XX
XX 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX Dean C, Gendall A;
XX WPI: 2000-499333/44.
XX N-PSDB; AAA47753.
XX
XX Isolated vernalization gene VRN2 is used to produce transgenic plants
XX with altered vernalization response, flowering time, leaf size and/or
XX shape or shade avoidance response for maximized reproductive success
XX
XX Claim 47; Page 71-72; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX plant encode polypeptides which are capable of affecting one or more
XX vernalization responses such as, flowering time, leaf size and/or
XX shape or the shade avoidance response of a plant into which the
XX nucleic acid is introduced. Introducing such sequences into plants
XX to alter these characteristics maximises the reproductive success of
XX the plant.
XX
XX Sequence 440 AA;
SQ
Query Match 96.1%; Score 2285; DB 21; Length 440;
Best Local Similarity 96.8%; Pred. No. 7.8e-225;
Matches 428; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
QY 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSFPLPCINRYKIGAKRR 60
DB 1 MCRONCRAKSSPEEVISTDENILTYCKPRLYNIFHLRSIGNSFPLPCINRYKIGAKRR 60
QY 61 KSRSTGMVYVFNKDCNNITLQKTEVREDSCSPFCSMGSGFKGLQFHLNSHDLFEFEFKL 120
DB 61 KSRSTGMVYVFNKDCNNITLQKTEVREDSCSPFCSMGSGFKGLQFHLNSHDLFEFEFKL 120
QY 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
QY 121 FEETQYVAVSVKINSFIPEEBSGDDKFEFSLCSKPKRRRGGRNNTRRLKYCFLPLD 180
DB 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHSSLDAGAKYILTSE 240
QY 181 SPSTLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHSSLDAGAKYILTSE 240
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
QY 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
DB 241 AVVPATKTRKLSAERSEARSHLLQKQFYHSHRVOPMALEQVMSDRSEDEVDVADF 300
QY 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
DB 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
QY 301 EDROMLDDFVDNKKDEKQFMHLMNSFVRKQRYIADGHSIMACEASRFEYKEELHRYSLF 360
DB 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420
QY 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420
DB 361 MCWRLEFLIKLMNHGLVDSATINNCAITLLENCRNSSDTTTNNNSVDRPSDNTNNNNIY 420

```

OY 421 DHPNDINNKNNVDKDNNSRDK 442
||||| |||||||||
Db 419 DHPNDIKNNKNNVDKDNNSRDK 440

RESULT 3
AAG54168
ID AAG54168 standard: Protein; 498 AA.
XX
AC AAG54168;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69039.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-01231825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139751.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151308.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158029.
PR 07-OCT-1999; 99US-0158232.
PR 08-OCT-1999; 99US-0158363.
PR 12-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160778.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161923.
PR 28-OCT-1999; 99US-0162142.
PR 29-OCT-1999; 99US-0162142.

Query Match 80.9%; Score 1923; DB 21; Length 498;
Best Local Similarity 89.4%; Fred. No. 1,le-187;
Matches 370; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

OY 43 PSLIPRLNKKIKAKRRKRSSTGVVFNKDCNNLTOKTEVREDSCSPFCSMTCGSFKG 102
DB 108 PSLIPRLNKKIKAKRRKRSSTGVVFNKDCNNLTOKTEVREDSCSPFCSMTCGSFKG 167
OY 103 LQHLNSSHDLFEFEFEKLEEVQTVNVSFKLNSFTFEERGSDDDKFEFSLCSRPKRKQ 162
DB 168 LQHLNSSHDLFEFEFEKLEEVQTVNVSFKLNSFTFEERGSDDDKFEFSLCSRPKRKQ 204

OY 163 RGRNNTRLKACFLPLDSPSLTNGTENGITLLDNGRSGTGYEATELIGOFEMTSNIP 222
DB 205 RGRNNTRLKACFLPLDSPSLTNGTENGITLLDNGRSGTGYEATELIGOFEMTSNIP 264
OY 223 ATASHSLDAGAVILTESAVVPARKTKTISAERSEAR-----SHILOKROFYH 271
DB 265 ATASHSLDAGAVILTESAVVPARKTKTISAERSEAR-----SHILOKROFYH 324
OY 272 SHRVOPMALEOVMSRDEDEVDVADFEEDROMLDFVYNKDEKQFPHLNSYRKOR 331
DB 325 SHRVOPMALEOVMSRDEDEVDVADFEEDROMLDFVYNKDEKQFPHLNSYRKOR 384
OY 332 VIADGHISMACFAFSREYKELHRYSSILFCWRLFLIKIMNGLVDSATINNCNTILENC 391
DB 385 VIADGHISMACFAFSREYKELHRYSSILFCWRLFLIKIMNGLVDSATINNCNTILENC 444
OY 392 RNSSDTTTNNNSVDPRSDSNTNNNIYDHPNDINNNKNNVDKNNNSRDKYIK 445
DB 445 RNSSDTTTNNNSVDPRSDSNTNNNIYDHPNDINNNKNNVDKNNNSRDKYIK 498
RESULT 4
AAC54169
ID MAG54169 standard; Protein; 367 AA.
XX
AC MAG54169;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 69040.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
FE 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126264.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131440.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136621.
PR 27-MAY-1999; 99US-0136692.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137528.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155186.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.

PR 26-OCT-1999; 990S-0161361.
PR 28-OCT-1999; 990S-0161920.
PR 28-OCT-1999; 990S-0161992.
PR 28-OCT-1999; 990S-0161993.
PR 29-OCT-1999; 990S-0162142.

Query Match 75.5%; Score 1796; DB 21; Length 367;
Best Local Similarity 88.7%; Pred. No. 6,8e-175;
Matches 346; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

OY 67 MYVFNKDCNNNTQKTEVREDSCSPSCMLGSPFGIDPFHNSHDLFEFPFKLFEFYOT 126
DB 1 MYVFNKDCNNNTQKTEVREDSCSPSCMLGSPFGIDPFHNSHDLFEFPFKLFEFYOT 56
OY 127 VNVSVKINSFTFEESGDDKFEPSLCSKPRKRRQGRNNTRKLKVCFLPLDSPSLTN 186
DB 57 VNPVIMFCS-----SKPRKRRQGRNNTRKLKVCFLPLDSPSLTN 97
OY 187 GTENGITLLDNGRGIGYPAPELACGFMTSNTIPPAIHAHSLDGAAYILTSEAAYPAT 246
DB 98 GTENGITLLDNGNGLGYPAPALAGCFMTSNTIPPAIHAHSLDGAAYILTSEAAYPAT 157
OY 247 KTKRLAERSEAR-----SHLLQKROFYTHSRVOPMALEQVMSDRDSEDEYDD 295
DB 158 KTKRLAERSEARSTALTEKISSHLLQKROFYTHSRVOPMALEQVMSDRDSEDEYDD 217
OY 296 DVADFEEDROMLDFVYNKDEKOFMHLANSFVRKORVIADGHIISWACEAFSFEYKELHR 355
DB 218 DVADFEEDROMLDFVYNKDEKOFMHLANSFVRKORVIADGHIISWACEAFSFEYKELHR 277
OY 356 YSSLFMCWRLFLIKILNNHGIYVSATINNCNTILENCRNSSDPTTTNNNSYDRPSDSTN 415
DB 278 YSSLFMCWRLFLIKILNNHGIYVSATINNCNTILENCRNSSDPTTTNNNSYDRPSDSTN 337
OY 416 NNNIVHPNDINNNKNNVNDKNNNRDKVIK 445
DB 338 NNNIVHPNDINNNKNNVNDKNNNRDKVIK 367

RESULT 5

AAG54170 ID AAG54170 standard; Protein; 339 AA.

AAG54170;

18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 69041.

KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PD 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 990S-0121825.

PR 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 23-MAR-1999; 990S-0126788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.
PR 23-APR-1999; 990S-0130891.
PR 28-APR-1999; 990S-0131449.
PR 30-APR-1999; 990S-0132048.
PR 30-APR-1999; 990S-0132407.
PR 04-MAY-1999; 990S-0132484.
PR 05-MAY-1999; 990S-0132485.
PR 06-MAY-1999; 990S-0132486.
PR 06-MAY-1999; 990S-0132487.
PR 07-MAY-1999; 990S-0132863.
PR 11-MAY-1999; 990S-0134256.
PR 14-MAY-1999; 990S-0134218.
PR 14-MAY-1999; 990S-0134219.
PR 14-MAY-1999; 990S-0134221.
PR 14-MAY-1999; 990S-0134370.
PR 14-MAY-1999; 990S-0134769.
PR 18-MAY-1999; 990S-0134941.
PR 20-MAY-1999; 990S-0135124.
PR 20-MAY-1999; 990S-0135353.
PR 24-MAY-1999; 990S-0135629.
PR 25-MAY-1999; 990S-0136021.
PR 27-MAY-1999; 990S-0136392.
PR 28-MAY-1999; 990S-0136782.
PR 01-JUN-1999; 990S-0137222.
PR 03-JUN-1999; 990S-0137528.
PR 04-JUN-1999; 990S-0137502.
PR 07-JUN-1999; 990S-0137724.
PR 08-JUN-1999; 990S-0138094.
PR 10-JUN-1999; 990S-0138540.
PR 14-JUN-1999; 990S-0138847.
PR 14-JUN-1999; 990S-0139119.
PR 16-JUN-1999; 990S-0139452.
PR 16-JUN-1999; 990S-0139453.
PR 17-JUN-1999; 990S-0139492.
PR 18-JUN-1999; 990S-0139454.
PR 18-JUN-1999; 990S-0139455.
PR 18-JUN-1999; 990S-0139456.
PR 18-JUN-1999; 990S-0139457.
PR 18-JUN-1999; 990S-0139458.
PR 18-JUN-1999; 990S-0139459.
PR 18-JUN-1999; 990S-0139460.
PR 18-JUN-1999; 990S-0139461.
PR 18-JUN-1999; 990S-0139462.
PR 18-JUN-1999; 990S-0139463.
PR 18-JUN-1999; 990S-0139750.
PR 18-JUN-1999; 990S-0138767.
PR 21-JUN-1999; 990S-0139817.
PR 22-JUN-1999; 990S-0139899.
PR 23-JUN-1999; 990S-0140353.
PR 23-JUN-1999; 990S-0140354.
PR 24-JUN-1999; 990S-0140695.
PR 28-JUN-1999; 990S-0140823.
PR 29-JUN-1999; 990S-0140991.
PR 30-JUN-1999; 990S-0141287.
PR 01-JUL-1999; 990S-0141842.
PR 01-JUL-1999; 990S-0142154.
PR 02-JUL-1999; 990S-0142055.
PR 06-JUL-1999; 990S-0142390.
PR 08-JUL-1999; 990S-0142803.
PR 09-JUL-1999; 990S-0142920.
PR 12-JUL-1999; 990S-0142977.
PR 13-JUL-1999; 990S-0143542.
PR 14-JUL-1999; 990S-0143624.
PR 15-JUL-1999; 990S-0144005.
PR 16-JUL-1999; 990S-0144089.
PR 16-JUL-1999; 990S-0144086.
PR 19-JUL-1999; 990S-0144325.
PR 19-JUL-1999; 990S-0144331.
PR 19-JUL-1999; 990S-0144332.
PR 19-JUL-1999; 990S-0144333.
PR 19-JUL-1999; 990S-0144334.
PR 19-JUL-1999; 990S-0144335.
PR 20-JUL-1999; 990S-0144332.

```

PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145210.
PR 23-JUL-1999; 99US-0145226.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147933.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159331.

```

```

PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 68.8%; Score 1635; DB 21; Length 339;
Best Local Similarity 87.8%; Pred. No. 1.8e-158;
Matches 318; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

OY 95 MLCGSEFKGLQFLHNSHDLFEPEFKLFEEYOTVNVSVKLSNFIIEEGSDDDKEPFSIC 154
DB 1 MLCGSEFKGLQFLHNSHDLFEPEFK-----YISVNPFTIMFCS----- 37
OY 155 SKPRRRRGGRNNRRRLKVCFLPLDPSLTNGTENGITLLDNGRGIGYEATLGAQF 214
DB 38 SKPRRRRGGRNNRRRLKVCFLPLDPSLTNGTENGITLLDNGRGIGYEATLGAQF 97
OY 215 EMTSNIPPAIHSSIDAGAKYILTSEAVVPATKTRKLSAERSEAR-----SHLL 263
DB 98 EMTSNIPPAIHSSIDAGAKYILTSEAVVPATKTRKLSAERSEARSTAILTEKISSHLL 157
OY 264 LÖKROFYSHRYVQPMALQVMSDRDSEDEVDVADFEEDROMLDPVDVNDKQFQHLW 323
DB 158 LÖKROFYSHRYVQPMALQVMSDRDSEDEVDVADFEEDROMLDPVDVNDKQFQHLW 217
OY 324 NSFVRKQVYADGHISWACEAFSRYEKELHRYSSLEWCWRLFLIKLNNHGLVDSATINN 383
DB 218 NSFVRKQVYADGHISWACEAFSRYEKELHRYSSLEWCWRLFLIKLNNHGLVDSATINN 277
OY 384 CNTILENCRNSSDTTNTNNNSVDPRSDSNTNNNNIYDHPNDINNNKNVNDKNNRDKV 443
DB 278 CNTILENCRNSSDTTNTNNNSVDPRSDSNTNNNNIYDHPNDINNNKNVNDKNNRDKV 337
OY 444 IK 445
DB 338 IK 339

RESULT 6
AAV53932
ID AAV53932 standard; Protein; 611 AA.
XX AAV53932;
AC
AC AAV53932;
DE
DE 13-MAR-2000 (first entry)
DE
DE A MPC1 protein having flowering regulating activity.
DE
DE Flowering regulating activity; MPC1; flowering; germination;
DE super early flowering mutation; altered flowering time;
DE flowering regulating gene; food crop; vegetable; flowering inhibition;
DE productivity.
DE
DE Arabidopsis thaliana.
OS
OS
XX
XX

```

FH Key Location/Qualifiers
 FT Domain 306..327
 FT /note="zinc finger domain"
 XX
 XX EP967278-A2.
 XX
 XX 29-DEC-1999.
 XX
 XX 28-JUN-1999; 99EP-0305077.
 XX
 XX 26-JUN-1998; 98JP-0180065.
 XX 24-JUN-1999; 99JP-0179043.
 XX
 XX (MITA) MITSUI CHEM INC.
 XX
 XX Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
 XX
 XX MPI: 2000-064612/06.
 XX N-PSDB: AA236947, AA236948.
 XX
 XX Novel DNA used to produce transgenic plants with altered floral
 XX regulation which can have increased crop yields -
 XX
 XX Claim 2, Page 12-15; 53pp; English.
 XX
 XX The present sequence represents a protein having a flowering regulating
 XX activity, which is designated MPC1. The genomic sequence is given
 XX in AA236948. A rice MPC1 is also disclosed in the specification. The
 XX rice and Arabidopsis cDNAs show significant homology with each other.
 XX A naturally occurring mutation of the MPC1 gene eliminates normal
 XX flowering regulating ability of plants, and leads to flowering
 XX immediately after germination (super early flowering mutation), and
 XX leads to a truncated protein comprising amino acids 1-540 of the present
 XX sequence. The MPC1 polynucleotide sequence can be used to produce plants
 XX with altered flowering times in comparison with wild type plants, by
 XX enhancing or inhibiting the expression of the flowering regulating gene.
 XX Antisense polynucleotides can be used to reproduce the effects of the
 XX mutated MPC1 gene. This alteration can be used to increase the yield
 XX of food crops. Flowering inhibition of vegetables increases their
 XX productivity.
 XX
 XX Sequence 611 AA:
 SQ
 Query Match 38.8%; Score 923; DB 21; Length 611;
 Best Local Similarity 34.8%; Pred. No. 3.5e-85;
 Matches 216; Conservative 52; Mismatches 108; Indels 244; Gaps 9;
 OY 1 MCQNCRAKSSPEVISTDENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLMYKIGARRR 60
 DB 1 MCHEDSRRLRISEEEIAEESLAAYCKPELVLYNIORRAIRNPFLORCLMYKIEAKHR 60
 OY 61 -----
 DB 61 RIOMTVFLSGAIDAGVOTOKLEPYLLIARLVSPKPAVSAVYRFSNACILIRGLGVDG 120
 OY 61 -----
 DB 61 VSOQAQNFLLPDMNRLALEAKSGSLAILFISPAQONSOQFIDSGKIHSGNIGCHLMSK 180
 OY 61 -----
 DB 61 -----
 OY 61 -----
 DB 181 IPIQSLYASWQKSPMNDIGQRYDVTSLVEMQPCFIKLKSKNEECVSIQVPSNLTSSSP 240
 OY 65 -----
 DB 241 QOVQVITISAEVGTSEKSPYSFSTNDISSSLQIIRLRGNVAFNRYRNNTLQKTEV 300
 OY 85 RECCGPFCSMLCGSEKGLQFHLNSHDLFEPEKLEFEYQTVAVSKLSFPEERGSD 144
 DB 301 TEDESCPFCLVCAASKGLRYHLPSTHDLNFEWVWEEOAVAVSLKTTMTSKVWEDD 360
 OY 145 -DKEFEPSLCSKP -KKRRORGRNNTRRLKVCFLPLDSPSLTG-----TENGITILAN 196
 OY 145 -DKEFEPSLCSKP -KKRRORGRNNTRRLKVCFLPLDSPSLTG-----TENGITILAN 196

DB 361 VDPKQOTFEFFSSKKRRRRRQKSGYSSRQ-----GPHLIGCEVLDKTDASHYS 411
 OY 197 DQNR---GLGPEATELAGOPEMTSNP-----PAIHSSLDGAKYILTSEAVPATKR 249
 DB 412 EKSRIPEKHY-----ERIGAESQGRVPETSPADYQSGDDPYGVASINGSTMLOFASTR 467
 OY 250 KISAERSRASHLLQKROFYHSHRQPMALBQVMSDRSEPDVDDVADFEEDRMLODF 309
 DB 468 KISIERSDLRNRSLQKROFPHSHRQPMALBQVMSDRSEPDVDDVADFEEDRMLODF 527
 OY 310 VDYNKEDEKOPMHLNNSFVRQRYIADGHISMACEASRYEKELHRYSSLPFCRFLIK 369
 DB 528 VDTKDEKQAMHNSFVRQRYIADGHISMACEASRYEKELHRYSSLPFCRFLIK 587
 OY 370 LMNGLVDSATINNCNTILE 389
 DB 588 LMNGLDARTMNCNTIFLE 607
 RESULT 7
 AAY53933
 ID AAY53933 standard; Protein: 604 AA.
 XX
 XX AAY53933;
 XX
 XX 13-MAR-2000 (first entry)
 XX
 XX An Os-MPC1 protein having flowering regulating activity.
 XX
 XX Flowering regulating activity; MPC1; flowering; germination;
 XX super early flowering mutation; altered flowering time;
 XX flowering regulating gene; food crop; vegetable; flowering inhibition;
 XX productivity.
 XX
 XX Oryza sativa.
 XX
 XX FH Key Location/Qualifiers
 FT Domain 310..335
 FT /note="zinc finger domain"
 XX
 XX EP967278-A2.
 XX
 XX 29-DEC-1999.
 XX
 XX 28-JUN-1999; 99EP-0305077.
 XX
 XX 26-JUN-1998; 98JP-0180065.
 XX 24-JUN-1999; 99JP-0179043.
 XX
 XX (MITA) MITSUI CHEM INC.
 XX
 XX Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
 XX
 XX MPI: 2000-064612/06.
 XX N-PSDB: AA236953.
 XX
 XX Novel DNA used to produce transgenic plants with altered floral
 XX regulation which can have increased crop yields -
 XX
 XX Claim 3, Page 36-39; 53pp; English.
 XX
 XX The present sequence represents a protein having a flowering regulating
 XX activity, which is designated Os-MPC1. An Arabidopsis MPC1 is also
 XX disclosed in the specification. The rice and Arabidopsis cDNAs show
 XX significant homology with each other. A naturally occurring mutation
 XX of the MPC1 gene eliminates normal flowering regulating ability of
 XX plants, and leads to flowering immediately after germination (super
 XX early flowering mutation). The MPC1 polynucleotide sequence can be
 XX used to produce plants with altered flowering times in comparison
 XX with wild type plants, by enhancing or inhibiting the expression of
 XX the flowering regulating gene. Antisense polynucleotides can be used
 XX to reproduce the effects of the mutated MPC1 gene. This alteration
 XX can be used to increase the yield of food crops. Flowering inhibition

CC of vegetables increases their productivity.
XX Sequence 604 AA;

Query Match 38.1%; Score 906.5; DB 21; Length 604;
Best Local Similarity 33.3%; Pred. No. 1.7e-83;
Matches 209; Conservative 59; Mismatches 109; Indels 251; Gaps 7;

```

QY 1 MCRONCRASSPEEVSTDENLLYCKPRLYNIIFHLRISGNSEFLPCLNTYIGAKRRR 60
   ||| ||| :||| :||| :||| ||| ||| :||| ||| ||| ||| ||| |||
DB 1 MCRHOPRRLSPDEOLAESEFALYCKPVELYNIORRSIKNPAFLQRCILYIKHARRK 60
QY 61 KSRST----- 65
   |||
DB 61 RSLITISLSGRTKELRAQNIPLVYLARPTNNVSLSEHSPIYRSRACLLTSFHEFG 120
   |||
QY 66 ----- 65
DB 121 KDYTEATFVLPVKNIATSRACSLNIIISCGRAEQTFDDNNGSGNHVEGSTLQLEKGC 180
QY 66 ----- 65
DB 181 FMGKIPIDILASSLGNVSLSGHTVEMSTVEMTPSELPKPLEDDSCLTFCGSRVDAT 240
QY 66 ----- -GMVFYNYKDCNNLTQ 80
   |||
DB 241 GSFQLOVSIQAQAGAKMSESPYSYVSYNDVPPSSLTHIRLRSGNVLENYKYNTNQ 300
QY 81 KTEVREDCSCPFCSMLGSGFKLQFHLNSHDLFEFEKLFEEYQTVNVSVKLSFIPE- 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 KTEVREDFSCPCLVPGSGFKIGCHLNASHDLFHYEFWISECQAVNVSLAKDTSWRTLE 360
QY 140 -EEGSDDDDEEPEPRLCSKPRKR-RQRGNNRRLKLVCLPLDSP-SLNGTENGTLIN 196
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 361 LAEGV-DPRHQTFSTYSRKRKRKVEISSKIRHVPHYDSSGPDQAQSGDDYQRE 419
QY 197 DGNRGLGYEATELACQFEMTSNIPALAHSSLDAGAKV---ILTSEAVVPATKTRKLSA 253
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 NGS-----SVAAHYDVPANSLHGSNLSA.PTVLQFCKTRKLS 456
QY 254 ERSEARSHLLQKQFYHSHRQPMALBOYMSDRDSEDEVDVAFEDROMLDDVDVA 313
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 457 ERADPPNRQLQKQFHFHRAQPMAMSKVFSDRDSEDEVDVADIAFEDRMLDDVDVA 516
QY 314 KDEKQFMHLMNSFVRKQRYIADGHISMACEAFGRFEKELHRSSLFMCWRLFLIKLWN 373
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 517 KDEKLIMHMNSFVRKQRYIADGHIPWACEAFSGFHGQELVQNPALIMCKRFPMVKLMNH 576
QY 374 GLVDSATINNCTIILENCRNSSDPTTNN 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 577 SLIDARAMNACNTILEGYLNGSSDPKKN 604

```

RESULT 8
AAB0062 standard; Protein; 107 AA.

XX AC AAB0062;
XX DT 16-NOV-2000 (first entry)
XX DE VRN2 polypeptide generated from aberrantly spliced VRN2 nucleic acid.
XX KM Vernalization gene; VRN2; plant characteristic; flowering time;
XX KW leaf size; leaf shape; shade avoidance response; reproduction;
XX KM breeding; pollination; cultivation.
XX OS Arabidopsis thaliana (Columbia).
XX PN WO200044918-A1.
XX PD 03-AUG-2000.
XX XX

PF 28-JAN-2000; 2000MO-GB00248.
XX PR 28-JAN-1999; 99GB-0001927.
XX PA (PLAN-) PLANT BIOSCIENCE LTD.
XX PI Dean C, Gendall A;
XX DR WPI: 2000-499333/44.
XX DR N-PSDB: AAA47755.
XX PT Isolated vernalization gene VRN2 is used to produce transgenic plants
XX PT with altered vernalization response, flowering time, leaf size and/or
XX PT shape or shade avoidance response for maximized reproductive success
XX PS Claim 47; Page 75; 105pp; English.
XX CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX CC plant encode polypeptides which are capable of affecting one or more
XX CC vernalization responses such as; flowering time, leaf size and/or
XX CC shape or the-shade avoidance response of a plant into which the
XX CC nucleic acid is introduced. Introducing such sequences into plants
XX CC to alter these characteristics maximises the reproductive success of
XX CC the plant. This polypeptide resulted from an aberrant splice in the
XX CC VRN2 gene.

SO Sequence 107 AA;
Query Match 23.1%; Score 549; DB 21; Length 107;
Best Local Similarity 99.0%; Pred. No. 4.9e-48;
Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MCRONCRASSPEEVSTDENLLYCKPRLYNIIFHLRISGNSEFLPCLNTYIGAKRRR 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MCRONCRASSPEEVSTDENLLYCKPRLYNIIFHLRISGNSEFLPCLNTYIGAKRRR 60
QY 61 KSRSTGMVFYNYKDCNNLTQKTEVREDCSCPFCSMLGSGSK 101
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 KSRSTGMVFYNYKDCNNLTQKTEVREDCSCPFCSMLGSGSK 101

```

RESULT 9
AAB01674 standard; protein; 813 AA.

XX AC AAB01674;
XX DT 01-SEP-2000 (first entry)
XX DE FIS2 protein sequence.
XX KM Seed; development; FIS; endosperm; autonomous embryogenesis;
XX KW transgenic plant; seedless fruit; parthenocarpic; citrus fruit;
XX KM stone fruit.
XX OS Arabidopsis thaliana.
XX FH Key location/Qualifiers
XX FT Misc-difference 1..120 /note="The first 120 residues are illegible in the
XX FT specification"
XX PN WO200016609-A1.
XX PD 30-MAR-2000.
XX PF 21-SEP-1999; 99WO-AU00805.
XX PR 21-SEP-1998; 98US-0101184.
XX PR 22-SEP-1998; 98AU-0006061.
XX PR 22-SEP-1998; 98AU-0006062.
XX PR 22-SEP-1998; 98AU-0006063.
XX PR 01-JUL-1999; 99AU-0001345.

YY	204	YPEATLQAGFEMTSNIPPAIAHSSLDGAKVILTSEAIVPATYTRKLSARSEARSHLL	263
YY	3	YPADE--CAQIVPENNLAP-----PAMIQFAKTRKLSIESRDMRNTL	43
YY	264	LQKQFTHSHRVCPMALEQVMSDDSEDEVDVDAFDEDRQMLDDFVDYVNDKQFHMILM	323
YY	44	LHKQFPHSHRAQPMALAEQVMSDDSEDEVDVDAFDEDRQMLDDFVDYVNDKQFHMILM	103
YY	324	NSFVR 328	
YY	104	NSFVR 108	
YY	RESULT 12		
YY	AAAG54171		
YY	ID	AAAG54171 standard; Protein: 63 AA.	
YY	AC	AAAG54171;	
YY	DT	18-OCT-2000 (first entry)	
YY	DE	Arabidopsis thaliana protein fragment SEQ ID NO: 69043.	
YY	KW	Protein identification; signal transduction pathway; metabolic pathway;	
YY	KM	hybridisation assay; genetic mapping; gene expression control; promoter;	
YY	KX	termination sequence.	
YY	OS	Arabidopsis thaliana.	
YY	PN	EP1033405-A2.	
YY	PD	06-SEP-2000.	
YY	PF	25-FEB-2000; 2000EP-0301439.	
YY	PR	25-FEB-1999; 99US-0121825.	
YY	PR	05-MAR-1999; 99US-0123180.	
YY	PR	09-MAR-1999; 99US-0123548.	
YY	PR	23-MAR-1999; 99US-0125788.	
YY	PR	25-MAR-1999; 99US-0126264.	
YY	PR	29-MAR-1999; 99US-0126785.	
YY	PR	01-APR-1999; 99US-0127462.	
YY	PR	06-APR-1999; 99US-0128234.	
YY	PR	08-APR-1999; 99US-0128714.	
YY	PR	16-APR-1999; 99US-0129845.	
YY	PR	19-APR-1999; 99US-0130077.	
YY	PR	21-APR-1999; 99US-0130449.	
YY	PR	23-APR-1999; 99US-0130510.	
YY	PR	23-APR-1999; 99US-0130891.	
YY	PR	28-APR-1999; 99US-0131449.	
YY	PR	30-APR-1999; 99US-0132048.	
YY	PR	30-APR-1999; 99US-0132407.	
YY	PR	04-MAY-1999; 99US-0132484.	
YY	PR	05-MAY-1999; 99US-0132485.	
YY	PR	06-MAY-1999; 99US-0132486.	
YY	PR	06-MAY-1999; 99US-0132487.	
YY	PR	07-MAY-1999; 99US-0132863.	
YY	PR	11-MAY-1999; 99US-0134256.	
YY	PR	14-MAY-1999; 99US-0134218.	
YY	PR	14-MAY-1999; 99US-0134219.	
YY	PR	14-MAY-1999; 99US-0134321.	
YY	PR	14-MAY-1999; 99US-0134370.	
YY	PR	18-MAY-1999; 99US-0134768.	
YY	PR	19-MAY-1999; 99US-0134941.	
YY	PR	20-MAY-1999; 99US-0135124.	
YY	PR	21-MAY-1999; 99US-0135353.	
YY	PR	24-MAY-1999; 99US-0135629.	
YY	PR	25-MAY-1999; 99US-0136021.	
YY	PR	27-MAY-1999; 99US-0136382.	
YY	PR	28-MAY-1999; 99US-0136782.	

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138887.
PR 16-JUN-1999; 99US-0139152.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 06-JUL-1999; 99US-0142055.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0143634.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144335.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145088.
PR 21-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 26-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145513.
PR 27-JUL-1999; 99US-0145518.
PR 27-JUL-1999; 99US-0145519.
PR 28-JUL-1999; 99US-0145591.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147302.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0152370.
PR 11-SEP-1999; 99US-0153708.
PR 12-SEP-1999; 99US-0153708.
PR 16-SEP-1999; 99US-0154038.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 29-SEP-1999; 99US-0156458.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157117.
PR 06-OCT-1999; 99US-0157553.
PR 07-OCT-1999; 99US-0157865.
PR 08-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.8%; Score 327; DB 21; Length 63;
Best Local Similarity 96.8%; Pred. No. 1,2e-25;
Matches 61; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 383 NCNTILENCRSSSTTTTNNNNNSYDRPSDSTNNNTVVDHPNDINKNNVNDKDNRSRDK 442
Db 1 NCNTILENWRNSDPTTNNNNNSYDRPSDSTNNNTVSHPNIDNNKNNVNDKDNRSRDK 60
|||
QY 443 VIK 445
|||
Db 61 VIK 63

RESULT 13
AAG61332
ID AAG61332 standard; Protein: 145 AA.
XX
AC AAG61332;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 79541.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PE 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139889.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149930.
 PR 23-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 26-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151903.
 PR 31-AUG-1999; 99US-0151938.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 04-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159295.
 PR 13-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 18-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 23-OCT-1999; 99US-0161404.
 PR 23-OCT-1999; 99US-0161405.
 PR 26-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161922.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 324; DB 21; Length 145;
 Best Local Similarity 98.3%; Pred. No. 8.3e-25;
 Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 43 PSELPRLNTYKIGAKRRKRSRSTGMVFNKDCNNTLQKTEVEDSCSPFCMTLCSGSK 101
 |||||||
 Db 81 PSELPRLNTYKIGAKRRKRSRSTGMVFNKDCNNTLQKTEVEDSCSPFCMTLCSGSK 139
 |||||||
 RESULT 14
 ID AAG61331 standard; Protein: 174 AA.
 AAG61331
 AC AAG61331;
 XX
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 79540.
 XX
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.
 OS Arabidopsis thaliana.
 XX
 PN EP103405-A2.
 PD
 XX
 PF 06-SEP-2000.
 XX
 XX 25-FEB-2000; 2000EP-0301439.
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131419.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134378.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136382.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137232.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137529.
 PR 07-JUN-1999; 99US-0137752.
 PR 08-JUN-1999; 99US-0138064.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0138847.
 PR 16-JUN-1999; 99US-0139419.
 PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140921.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149375.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149729.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 13.6%; Score 324; DB 21; Length 174;
Best Local Similarity 98.3%; Pred. No. 1, 1e-24;
Matches 58; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 43 PSFLPRCLNFKIGARRKRSRSTGMVFNFKDCNNNTLOKTEVREDCSCPFCSMLGSPK 101
|||||
Db 110 PSFLPRCLNFKIGARRKRSRSTGMVFNFKDCNNNTLOKTEVREDCSCPFCSMLGSPK 168

RESULT 15

```

AAB00063
ID AAB00063 standard; Protein; 154 AA.
XX
AC AAB00063;
XX
DT 16-NOV-2000 (first entry)
XX
DE C72616 Expressed sequence tag (modified).
XX
XX Vernalization gene: VRN2; plant characteristic: flowering time;
KM leaf size; leaf shape; shade avoidance response; reproduction;
XX breeding; pollination; cultivation.
XX
OS Unspecified.
XX
XX WO200044918-A1.
XX
PD 03-AUG-2000.
XX
FE 28-JAN-2000; 2000WO-GB00248.
XX
PR 28-JAN-1999; 99GB-0001927.
XX
PA (PLAN-) PLANT BIOSCIENCE LTD.
XX
PI Dean C, Gendall A;
XX
DR WPI; 2000-499333/44.
XX
DR N-PSDB; AAA47756.
XX
PT Isolated vernalization gene VRN2 is used to produce transgenic plants
PT with altered vernalization response, flowering time, leaf size and/or
PT shape or shade avoidance response for maximized reproductive success
XX
PS Disclosure; Page 76; 105pp; English.
XX
XX Isolated nucleic acid sequences obtained from the VRN2 locus of a
CC plant encode polypeptides which are capable of affecting one or more
CC vernalization responses such as, flowering time, leaf size and/or
CC shape or the shade avoidance response of a plant into which the
CC nucleic acid is introduced. Introducing such sequences into plants
CC to alter these characteristics maximises the reproductive success of
CC the plant. This cDNA was generated from an aberrantly spliced VRN2
XX gene.
XX
SQ Sequence 154 AA;
XX
Query Match 13.1%; Score 311; DB 21; Length 154;
Best Local Similarity 62.1%; Pred. No. 2e-23;
Matches 64; Conservative 14; Mismatches 21; Indels 4; Gaps 2;
OY 223 AIAHSGLDAGAKV--ILTSEAVVATYTRKLSERSEARSHLLQKRFYHSHRVOPMA 279
DB 53 SVAHASVDPANSLHGSNLSAPTVLQFGKTRKLSVERADPRNRQLQKRFHSHRAQPMMA 112
OY 280 LEQVMSDRDSEDEVDDVADPFEDROMLDDFVDVANKDEKQEFMHL 322
DB 113 LGAVFSDRDSDEVDVDDIADFEEDROMLDDFVDVANKDE-LIMHM 154

```

Search completed: June 13, 2003, 15:35:41
 Job time : 74 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:37:19 ; Search time 26 Seconds

(without alignments)
503.585 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRACKSPPEVISTDE.....INNNNNVDKNDSKVIK 445

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

262574

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA.*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep.*

2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep.*

3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep.*

4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep.*

5: /cgn2_6/ptodata/1/1aa/PCRTUS.COMB.pep.*

6: /cgn2_6/ptodata/1/1aa/Backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126.5	5.3	1584	US-09-457-040B-27	Sequence 27, Appl
2	123	5.2	110	US-08-569-165-34	Sequence 34, Appl
3	112	4.7	361	US-08-415-751-3	Sequence 3, Appl
4	105	4.4	732	US-08-914-999-8	Sequence 8, Appl
5	103	4.3	432	US-08-700-152A-4	Sequence 4, Appl
6	103	4.3	1073	US-09-541-782-6	Sequence 6, Appl
7	103	4.3	1073	US-09-723-820-6	Sequence 6, Appl
8	98.5	4.1	1388	US-09-572-191-2	Sequence 2, Appl
9	98.5	4.1	1388	US-09-723-262-2	Sequence 2, Appl
10	98.5	4.1	1388	US-09-723-219-2	Sequence 2, Appl
11	98	4.1	652	US-08-559-896B-2	Sequence 2, Appl
12	94	4.0	2391	US-08-446-855A-2	Sequence 2, Appl
13	94	4.0	2391	US-09-150-741-2	Sequence 2, Appl
14	92	3.9	286	US-09-384-162-6	Sequence 2, Appl
15	91.5	3.8	689	US-09-177-249-2	Sequence 2, Appl
16	91.5	3.8	689	US-09-061-769A-2	Sequence 2, Appl
17	91.5	3.8	832	US-08-677-734A-12	Sequence 12, Appl
18	91.5	3.8	832	US-09-097-053-12	Sequence 12, Appl
19	91.5	3.8	1786	US-08-973-462-8	Sequence 8, Appl
20	91	3.8	1051	US-08-965-762-14	Sequence 14, Appl
21	91	3.8	1051	US-09-911-927-14	Sequence 14, Appl
22	91	3.8	1051	US-09-911-882-14	Sequence 14, Appl
23	90.5	3.8	1099	US-09-442-100-2	Sequence 2, Appl
24	90.5	3.8	428	US-09-118-319-5	Sequence 5, Appl
25	89.5	3.8	581	US-08-619-812-6	Sequence 6, Appl
26	89.5	3.8	826	US-07-638-431-2	Sequence 2, Appl
27	89.5	3.8	826	PCT-US92-00018-2	Sequence 2, Appl

28	89	3.7	749	2	US-08-568-459A-6	Sequence 6, Appl
29	89	3.7	749	2	US-08-487-826B-6	Sequence 6, Appl
30	89	3.7	749	4	US-09-210-288-6	Sequence 6, Appl
31	88.5	3.7	432	3	US-09-118-319-2	Sequence 2, Appl
32	88.5	3.7	507	4	US-09-134-001C-3978	Sequence 3978, Ap
33	88.5	3.7	596	4	US-09-752-165-2	Sequence 2, Appl
34	88.5	3.7	911	3	US-09-074-579-4	Sequence 4, Appl
35	88.5	3.7	911	3	US-09-388-774-4	Sequence 4, Appl
36	88	3.7	930	4	US-09-134-001C-5314	Sequence 5314, Ap
37	87.5	3.7	392	3	US-08-523-373-6	Sequence 6, Appl
38	87.5	3.7	1081	2	US-08-843-530B-18	Sequence 18, Appl
39	87.5	3.7	1117	2	US-08-843-530B-33	Sequence 33, Appl
40	87	3.7	620	1	US-08-484-105-4	Sequence 4, Appl
41	87	3.7	620	1	US-08-484-105-4	Sequence 4, Appl
42	87	3.7	10182	4	US-09-134-001C-3159	Sequence 13, Appl
43	86.5	3.6	770	2	US-08-209-521-13	Sequence 123, App
44	86.5	3.6	770	4	US-08-961-810-123	Sequence 123, App
45	86.5	3.6	770	4	US-08-352-902D-123	Sequence 123, App

ALIGNMENTS

RESULT 1	US-09-457-040B-27	Sequence 27, Application US/09457040B
Patent No. 6387641		
GENERAL INFORMATION:		
APPLICANT: Vertex Pharmaceuticals Incorporated		
APPLICANT: Bellon, Steve		
TITLE OF INVENTION: Crystallized P38 Complexes		
FILE REFERENCE: VPI/98-14		
CURRENT APPLICATION NUMBER: US/09/457,040B		
CURRENT FILING DATE: 1999-12-08		
NUMBER OF SEQ ID NOS: 41		
SOFTWARE: PatentIn version 3.0		
SEQ ID NO 27		
LENGTH: 1584		
TYPE: PRT		
ORGANISM: DICDI - Dictyostellium Discoideum		
US-09-457-040B-27		
Query Match	5.3%; Score 126.5; DB 4; Length 1584;	
Best Local Similarity	22.2%; Pred. No. 0.00056;	
Matches	77; Conservative 37; Mismatches 100; Indels 133; Gaps 16;	
QY	149 EEFSLCKRKRKRGGRNTRRLKVCPLDPSLFTNGTITLNDGNGLGPEAT	208
DB	210 EEEFVCRMSKKNYH-GKHVYVNRKTAFLDPSKDTIG--GVRATOPFGGFCF----	261
QY	209 ELAOGFEMTSNIPAIHSSLDGAKVILTSEAVVPATRTKLSAERSEARSHLLLOKRO	268
DB	262 -----FE-----VIIDQLKGQSLISLANLEVPYF-----YHGVMPRS	295
QY	269 F-YHSRHPQALAEQMSDRDSEDEYDDVADFEEDROMLDDVYVKNDEKQTMHLNMFV	327
DB	296 YGVH-----NDDGRFRFMEEE-----PGVNEGSS-----YGSYSY	324
QY	328 RKQRYADGSHSMACE-----AFSFRFEKELHRSYSLFWCQRLF-----	366
DB	325 KKGDIIGCG-LSTFSREIRFTTNGMYLGTAFSNVT-----GVYPSVANEPEGISIT	375
QY	367 -----LRLKLNHGLVDSATI-----NMCNTILECRNSSDPT	398
DB	376 GVYGFPRKRSQVTLMLKN---VNSGSLVLPVNNNNNNNNNNNNNNNNNNNNNNNNNNNN	432
QY	399 TT-----NNNSVDRPSDSTNNNNIYDHPDINKNNKNDKNNNS	439
DB	433 TSTSPSSINNEDISSNN	479
RESULT 2	US-08-569-166-34	

Fri Jun 20 09:03:39 2003

us-09-890-220-2.ra1

Page 2

```
? Sequence 34, Application US/08569166
? Patent No. 5830722
? GENERAL INFORMATION:
? APPLICANT: NICOLAS, LUC
? APPLICANT: CHARLES, JEAN-FRANCOIS
? APPLICANT: DELECLUSE, ARMELE
? APPLICANT: BARLOY, FREDERIQUE
? TITLE OF INVENTION: CLOSTRIDIUM BIFERMENTANS DNA FRAGMENT
? TITLE OF INVENTION: BEARING GENES CODING FOR PROTEINS LINKED TO AN
? TITLE OF INVENTION: INSECTICIDAL ACTIVITY
? NUMBER OF SEQUENCES: 35
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
? STREET: 1755 S. JEFFERSON DAVIS HIGHWAY, SUITE 400
? CITY: ARLINGTON
? STATE: VA
? COUNTRY: USA
? ZIP: 22202
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/569,166
? FILING DATE: 05-JUL-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: PCT/FR94/00768
? FILING DATE: 24-JUN-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: FR 93/07795
? FILING DATE: 25-JUN-1993
? ATTORNEY/AGENT INFORMATION:
? NAME: OBLON, NORMAN F.
? REGISTRATION NUMBER: 24,618
? REFERENCE/DOCKET NUMBER: 660-106-0 PCT
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 703-413-2220
? TELEFAX: 703-413-3000
? INFORMATION FOR SEQ ID NO: 34:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 110 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? US-08-569-166-34

Query Match
Best Local Similarity 5.2%; Score 123; DB 2; Length 110;
Matches 24; Conservative 16; Mismatches 18; Indels 4; Gaps 1;
QY 380 TNNCNTILCNCNSSDTTNNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVYDKDNNS 439
Db 10 TNNNTNNNNNNNNNNNNNNNNNA---NNQNTNNNNNNNNNNNNNNNNNNNNHQQNNNN 65
```

```
QY 440 RD 441
Db 66 NN 67

RESULT 3
Sequence 3, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LERCH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
```

```
? TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
? TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
? TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
? TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
? NUMBER OF SEQUENCES: 50
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: PHILLIPS, MOORE, LEMPIO & FINLEY
? STREET: 385 Sherman Avenue, Suite 6
? CITY: Palo Alto
? STATE: California
? COUNTRY: United States of America
? ZIP: 94306-1840
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
? SOFTWARE: Wordperfect 5.1
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/415,751
? FILING DATE: 03-APR-1993
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/071,880
? FILING DATE: June 1, 1993
? APPLICATION NUMBER: 07/891,301
? FILING DATE: May 29, 1992
? ATTORNEY/AGENT INFORMATION:
? NAME: Hana Dolezalova
? REGISTRATION NUMBER: 30,518
? REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (415) 324-1677
? TELEFAX: (415) 324-1678
? INFORMATION FOR SEQ ID NO: 3:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 361 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? ORGANISM: Cryptosporidium parvum
? FEATURE:
? NAME/KEY: Positions coded by nonsense codons are
? US-08-415-751-3

Query Match
Best Local Similarity 4.7%; Score 112; DB 1; Length 361;
Matches 23; Conservative 12; Mismatches 25; Indels 0; Gaps 0;
QY 382 NNCNTILCNCNSSDTTNNNNNSVDRPSDSNTNNNNIVDHPNDINNKNNVYDKDNNSD 441
Db 115 NNHNNSXSNHNNNSKTNNNNNNGSHXATNSHNNNNIKANSYNNNNKSNNNNNNSAND 174
```

```
QY 4
Db 914-999-8
Sequence 8, Application US/08914999
Patent No. 6346406
GENERAL INFORMATION:
APPLICANT: Ryazanov, Alexey G.
APPLICANT: Pavut, Karen S.
APPLICANT: Hailt, William N.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
TITLE OF INVENTION: AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
STREET: Floor
CITY: Hackensack
```

```

STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostellium discoideum
US-08-914-999-8

```

Query Match 4.4%; Score 105; DB 4; Length 732;

Best Local Similarity 20.1%; Pred. No. 0.033; Indels 74; Gaps 17;

```

Matches 76; Conservative 61; Mismatches 167;
QY 117 EFKLE-EXQIVNVSVKLNSFLFEDEG---SDDKFEPPSLCSKPRKRORGRNTRRL 172
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 34 EKKIRFPYGTESQTLQIGKMLPSGGGATADSKFEKFKARNTLADIQYKVDLYVAV 93
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 173 KV-----CFLPLDSPSLNTEGNTITLNDGNRLGYPEATELAGOFEMTSNI-----PP 222
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 94 KSKRTNDSLPTLTLIAFLDGSERAIKWEYD-----PYTF--TAQWTCITLVKVEPV 144
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 223 AIAHSLDAGAKVILTSEAVVPATKRLSAERSEARSHLLQKROF-----YHSHR 274
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 145 PPAEAFKRAYHTLDSKSGAGRYVSKIGKKPTFRSPFEDYKQMIKAKKADYNSRK 204
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 215 VOPMALEOVMS-----DRDSEDEV--DDDVADFEEDROMLDFVYNKDEK---QFMHL 322
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 205 -PPKRIEFLQSCVLEFVDRTSDDLICGAEPYVEGQYRKYNNSGFSVNDERNTPOSFHF 263
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 323 WNSFVRKORVLTAD-----GHISWACFAFSHFYEKEKELHRYSSLFCWCRFLIKLMNGIVDS 378
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 264 TYEHSNHOLLIDIDGVC-----DHYTDPQIHYYDGV---GFGIGNLGQKGEFEK 310
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 379 ATINNCCNTILE--NCRNSDPTTTNNNSVDRPS-----DSNTNNNTIVDHPNDIN 427
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 311 LDTHRCMAICQYLNQISINPKSEKSDCGVPRPDLIFPTISRDNNNNNNN--NNNNNNN 368
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 428 KKNANVDKDNNSRDKVIK 445
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 369 NNNNNNNNNNNNNSSISK 386
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 5

US-08-700-152A-4

; Sequence 4, Application US/08/00152A

; Patent No. 5994622

; GENERAL INFORMATION:

; APPLICANT: Jofuku, K. Diane

; APPLICANT: Okamura, Jack K.

; TITLE OF INVENTION: Methods for Improving Seeds

; NUMBER OF SEQUENCES: 4

```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/700,152A
FILING DATE: 20-AUG-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-067200US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 432 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-700-152A-4

```

Query Match 4.3%; Score 103; DB 2; Length 432;

Best Local Similarity 21.5%; Pred. No. 0.023; Indels 134; Gaps 20;

```

Matches 88; Conservative 54; Mismatches 134;
QY 74 DCNNTLOKTEVREDCSCPC-----SMLGSPFKLOFHNLSSHDLEFEFKLEEXQIVNV 129
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 DLNDAPHQHQ-REESEEECCYSSPKRVGSF-----NSS----- 36
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 130 SVKLSFLFEESGSDDDKFEPPSLCSKPRKRORGRNTRRLKACFLPLDSPSLNTE 189
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 37 ---SSAVVIEDGSDDEL-----NRRAPNPVLTGHOFPEMDS----- 71
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 190 NGITLNDGNRLGYPEATELAGOFEMTSNIPPATAHSLDAGAKVILTSEAVVPATKTR 249
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 72 -----NGGCVASGFPRAHMFVGFQCSLALGSSAGKATVNAAV---EAGQPLKSR 122
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 250 KLSAERSEA-----RSHLLQKROFY-----HSHRVOPMALEOVMSDRDSED 291
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 123 RPRRSRSQYRGVTFEYRTGRWESHIMDCGQVYLGEPDTAAHAAAYDRAAIKFRGVEA 182
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 292 EYDDVDVADDED--ROMLDFVYNKDEQOFMLMNSFVVKORVIVADGHISWACFAFSRY 349
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 183 DTFENIDVDDDLKQK-----TNLTKEEFVHV---LKRQST-----GPPRS 221
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 350 EKELRHYSLSFCWCR-----LFLIKLMNH--GLVDSATINNCCNTILENCR-----N 393
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 222 SK--YRGVTLHKCGWEARMGQFLGKKYYLLDF-----TEVEARAYDKAAIKCN 271
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 394 SSDTTTNNNSVDRP--SDSNTNNNTIVDHPNDINNNKNNVNNKDNNSD 441
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 272 GDAVATNFDPSIYDEELNASSGNPTTPQDHMLDLISLGNSSANSK-HKSQD 320
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

RESULT 6

US-09-541-782-6

; Sequence 6, Application US/09/541782

; Patent No. 6284480

; GENERAL INFORMATION:

; APPLICANT: Nislow, Corey

; APPLICANT: Sakowicz, Roman

; APPLICANT: Beraud, Christophe

; TITLE OF INVENTION: Antifungal Assay

Fri Jun 20 09:03:39 2003

us-09-890-220-2.ra1

Page 4

```
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/541.782
CURRENT FILING DATE: 2000-04-03
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1073
TYPE: PRT
ORGANISM: S.pombe
US-09-541-782-6
```

```
Query Match
Best Local Similarity 18.3%; Score 103; DB 4; Length 1073;
Matches 108; Conservative 75; Mismatches 200; Indels 206; Gaps 24;
```

```
QY 13 EEVISTDE-----NLIYCKPVRLYNIFHLRSIGNPSFLPRCLNYKIGAKR 58
DB 394 EETISTLEYAARAKSIRKPNQNLVFRKVLKIDLVDIRLKN-----LN-----A 441
QY 59 KRKRSSTGVVYVYKDCNNLTOKTEVREDCSCPCSMGSGFKGLQFHLNSHDLFEFEF 118
DB 442 TRKKNVYLAESTYKELMDRYQNKDL-----LCQEQARKLEVLDLVANKSSRQLQYVS 494
QY 119 KLPEEY-----QTVNVSVKLNSEIFEEGSDDD-----KEEP-----150
DB 495 KSNQEHKEVEALQDLVNSSTELSVKSENEKLNELVLEIKRKKEYTEAKITYAT 554
QY 151 -----FSLCK--PKRRRQGRNNTRRLKYCFPL--DSPSLTNGTNGCI 192
DB 555 DLSQYRESKEVIALSYEKLDRTERRNNKENENNFWNKFNLLTMLRSFGHSFTDENGIF 614
QY 193 TLNDGNRGIG-----YPEATLAGQFE-----MTSNIPPAIAHSSLDGAKVI 236
DB 615 TLNDPNSMEBLNTHSNOLLISMTKITEHFOSLDEALQASRSCAVNSSL-----L 669
QY 237 LTSE-----AVPA-----TKTKLSAERSEARSHLLQKROFY 270
DB 670 IVELKDSKNSLIDALEHSLQDISMSQKLGNGISSELELOKDKESYROLVDELRSLY 729
QY 271 ---HSH-RVQPMALAEQVMSDRS-----EDEVDVADFEQROMLDFV 310
DB 730 NLOHTHEESQKELMYGVANDIDALVKTCTTSLNDADITLSDYISQKSFESKOO-DLIA 788
QY 311 DVNDEKOFM-----HL--WNSFVKORVIADG-----336
DB 789 NIGKIYVSNFLOEONESLYTKADILHSLHNTNSIRKANEIMNNSEEFELRNASQAIEV 848
QY 337 ---HSMACEAFSRFE---KELHRYSSLFCWRLFLIKLMNHGLVDSATINNCTIL 388
DB 849 GANKERIOKTVENGSQLDSKSKAIHSNRSRMYDCHLALAESQKGV--NLEVOTLDRLL 906
QY 389 ENCRRNSDPTT---NNNSVDRPSDSNTNNNNIVDH 422
DB 907 QVKHESDNTKKEHQQLDLLESLVGNNDNLDSIKTPHTELQKTIKH 955
```

```
RESULT 7
US-09-723-820-6
Sequence 6, Application US/09723820
Patent No. 6468760
GENERAL INFORMATION:
APPLICANT: Nislow, Corey
APPLICANT: Beraud, Christophe
TITLE OF INVENTION: Antitungal Assay
FILE REFERENCE: 1015
CURRENT APPLICATION NUMBER: US/09/723,820
CURRENT FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 09/541,782
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
```

```
LENGTH: 1073
TYPE: PRT
ORGANISM: S.pombe
US-09-723-820-6
```

```
Query Match
Best Local Similarity 18.3%; Score 103; DB 4; Length 1073;
Matches 108; Conservative 75; Mismatches 200; Indels 206; Gaps 24;
```

```
QY 13 EEVISTDE-----NLIYCKPVRLYNIFHLRSIGNPSFLPRCLNYKIGAKR 58
DB 394 EETISTLEYAARAKSIRKPNQNLVFRKVLKIDLVDIRLKN-----LN-----A 441
QY 59 KRKRSSTGVVYVYKDCNNLTOKTEVREDCSCPCSMGSGFKGLQFHLNSHDLFEFEF 118
DB 442 TRKKNVYLAESTYKELMDRYQNKDL-----LCQEQARKLEVLDLVANKSSRQLQYVS 494
QY 119 KLPEEY-----QTVNVSVKLNSEIFEEGSDDD-----KEEP-----150
DB 495 KSNQEHKEVEALQDLVNSSTELSVKSENEKLNELVLEIKRKKEYTEAKITYAT 554
QY 151 -----FSLCK--PKRRRQGRNNTRRLKYCFPL--DSPSLTNGTNGCI 192
DB 555 DLSQYRESKEVIALSYEKLDRTERRNNKENENNFWNKFNLLTMLRSFGHSFTDENGIF 614
QY 193 TLNDGNRGIG-----YPEATLAGQFE-----MTSNIPPAIAHSSLDGAKVI 236
DB 615 TLNDPNSMEBLNTHSNOLLISMTKITEHFOSLDEALQASRSCAVNSSL-----L 669
QY 237 LTSE-----AVPA-----TKTKLSAERSEARSHLLQKROFY 270
DB 670 IVELKDSKNSLIDALEHSLQDISMSQKLGNGISSELELOKDKESYROLVDELRSLY 729
QY 271 ---HSH-RVQPMALAEQVMSDRS-----EDEVDVADFEQROMLDFV 310
DB 730 NLOHTHEESQKELMYGVANDIDALVKTCTTSLNDADITLSDYISQKSFESKOO-DLIA 788
QY 311 DVNDEKOFM-----HL--WNSFVKORVIADG-----336
DB 789 NIGKIYVSNFLOEONESLYTKADILHSLHNTNSIRKANEIMNNSEEFELRNASQAIEV 848
QY 337 ---HSMACEAFSRFE---KELHRYSSLFCWRLFLIKLMNHGLVDSATINNCTIL 388
DB 849 GANKERIOKTVENGSQLDSKSKAIHSNRSRMYDCHLALAESQKGV--NLEVOTLDRLL 906
QY 389 ENCRRNSDPTT---NNNSVDRPSDSNTNNNNIVDH 422
DB 907 QVKHESDNTKKEHQQLDLLESLVGNNDNLDSIKTPHTELQKTIKH 955
```

```
RESULT 8
US-09-572-191-2
Sequence 2, Application US/09572191
Patent No. 6355466
GENERAL INFORMATION:
APPLICANT: Beraud, Christophe
APPLICANT: Sakowicz, Roman
TITLE OF INVENTION: No. 6355466 motor proteins and methods for
FILE REFERENCE: 1017
CURRENT APPLICATION NUMBER: US/09/572,191
CURRENT FILING DATE: 2000-05-17
NUMBER OF SEQ ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 1388
TYPE: PRT
ORGANISM: Human
US-09-572-191-2
```

```
Query Match
Best Local Similarity 19.0%; Score 98.5; DB 4; Length 1388;
```


Matches 79; Conservative 81; Mismatches 150; Indels 105; Gaps 18;

QY 85 REDCGCPFCMGLGSGFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLSNF----- 136
 Db 573 KEPCLEFANTKLAQQLQIOTELNNSKQYE-EFKELTRKQOLESELSQOKANLNE 631
 QY 137 -IFE-----EGSDDDKF--EPFSLCKPRKRGGRNNTRLKVCFLPLDPSLTN 186
 Db 632 NLEATKACKROEVSQLNKIHAETLKITTPRKAYQLHSRP-----VKLSPEM-- 680
 QY 187 GTENGITLLDNGNRGLGYPEATELAGQFEMTSNIPPAIHAHSSLDGAKYILSEAVVPAT 246
 Db 681 -----GSFSLYQNSSIILDNDILNEPVPENNEQAFEAISELRTVQEQMSAL 729
 QY 247 KTRKLSAERSEARSHLLQK---ROFYHSHRVQPMALQVMSDRSEDEVDV---AD 299
 Db 730 QA-KLD---EEHKNKLKQOHVCKLEHSTQMOELFSSRIDWTQOEELLSQANVLEKQ 785
 QY 300 FEDROMLDPDVYNKDEKQFMHLMNSFYVKQRYI---ADGHISMACEAFSRF---YEKEL 353
 Db 766 LOETQTKNDFL-----KSEVHDLRVVLHSAADKELSVKLEYSSFKTNOKEKF 832
 QY 354 HRYSSLFWCWRFL--IKLMHGLVDSAT-----INNCTILLECRNSS 395
 Db 833 NKLSEHMHVQQLDNLRLNEKLESKACLODSYDNLQEIKKFEIDQLSRNLQNFKEEN 892
 QY 396 DTTTNNNNNSVDRPSDSTNNNNIYDHPDINNKNV---DNKDNNSRD--KVIK 445
 Db 893 ETLKSDLNLMF-----LLEAEKERNNKLSLOPEEDKENSKEILKYLE 936

RESULT 9
 US-09-723-262-2
 ; Sequence 2, Application US/09723262
 ; Patent No. 6379912
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6379912el motor proteins and methods for
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723,262
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 09/572,191
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-723-262-2

Query Match 4.1%; Score 98.5; DB 4; Length 1388;
 Best Local Similarity 19.0%; Pred. No. 0.46;
 Matches 79; Conservative 81; Mismatches 150; Indels 105; Gaps 18;

QY 300 FEDROMLDPDVYNKDEKQFMHLMNSFYVKQRYI---ADGHISMACEAFSRF---YEKEL 353
 Db 766 LOETQTKNDFL-----KSEVHDLRVVLHSAADKELSVKLEYSSFKTNOKEKF 832
 QY 354 HRYSSLFWCWRFL--IKLMHGLVDSAT-----INNCTILLECRNSS 395
 Db 833 NKLSEHMHVQQLDNLRLNEKLESKACLODSYDNLQEIKKFEIDQLSRNLQNFKEEN 892
 QY 396 DTTTNNNNNSVDRPSDSTNNNNIYDHPDINNKNV---DNKDNNSRD--KVIK 445
 Db 893 ETLKSDLNLMF-----LLEAEKERNNKLSLOPEEDKENSKEILKYLE 936

RESULT 10
 US-09-723-219-2
 ; Sequence 2, Application US/09723219
 ; Patent No. 6391613
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Wood, Kenneth
 ; TITLE OF INVENTION: No. 6391613el motor proteins and methods for
 ; FILE REFERENCE: 1017
 ; CURRENT APPLICATION NUMBER: US/09/723,219
 ; CURRENT FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: US 09/572,191
 ; PRIOR FILING DATE: 2000-05-17
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 1388
 ; TYPE: PRT
 ; ORGANISM: Human
 US-09-723-219-2

Query Match 4.1%; Score 98.5; DB 4; Length 1388;
 Best Local Similarity 19.0%; Pred. No. 0.46;
 Matches 79; Conservative 81; Mismatches 150; Indels 105; Gaps 18;

QY 85 REDCGCPFCMGLGSGFKGLQFHLNSSHDLFEFEFKLFEEYQTVNVSVKLSNF----- 136
 Db 573 KEPCLEFANTKLAQQLQIOTELNNSKQYE-EFKELTRKQOLESELSQOKANLNE 631
 QY 137 -IFE-----EGSDDDKF--EPFSLCKPRKRGGRNNTRLKVCFLPLDPSLTN 186
 Db 632 NLEATKACKROEVSQLNKIHAETLKITTPRKAYQLHSRP-----VKLSPEM-- 680
 QY 187 GTENGITLLDNGNRGLGYPEATELAGQFEMTSNIPPAIHAHSSLDGAKYILSEAVVPAT 246
 Db 681 -----GSFSLYQNSSIILDNDILNEPVPENNEQAFEAISELRTVQEQMSAL 729
 QY 247 KTRKLSAERSEARSHLLQK---ROFYHSHRVQPMALQVMSDRSEDEVDV---AD 299
 Db 730 QA-KLD---EEHKNKLKQOHVCKLEHSTQMOELFSSRIDWTQOEELLSQANVLEKQ 785
 QY 300 FEDROMLDPDVYNKDEKQFMHLMNSFYVKQRYI---ADGHISMACEAFSRF---YEKEL 353
 Db 766 LOETQTKNDFL-----KSEVHDLRVVLHSAADKELSVKLEYSSFKTNOKEKF 832
 QY 354 HRYSSLFWCWRFL--IKLMHGLVDSAT-----INNCTILLECRNSS 395
 Db 833 NKLSEHMHVQQLDNLRLNEKLESKACLODSYDNLQEIKKFEIDQLSRNLQNFKEEN 892
 QY 396 DTTTNNNNNSVDRPSDSTNNNNIYDHPDINNKNV---DNKDNNSRD--KVIK 445
 Db 893 ETLKSDLNLMF-----LLEAEKERNNKLSLOPEEDKENSKEILKYLE 936

RESULT 11
 US-08-559-896B-2
 ; Sequence 2, Application US/08559896B

Fri Jun 20 09:03:39 2003

us-09-890-220-2.ra1

Page 6

```
; Patent No. 6310046
; GENERAL INFORMATION:
; APPLICANT: Patrick E. Duffy
; APPLICANT: Christian F. Ockenhause
; TITLE OF INVENTION: SEQUESTRAIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John Moran
; STREET: USA MRMC - MCMR-JA
; CITY: FORT DETRICK, FREDERICK
; STATE: MARYLAND
; COUNTRY: USA
; ZIP: 21702-5012
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/559,896B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Moran, John
; REGISTRATION NUMBER: 26,313
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 619-2065
; TELEFAX: (301) 619-7714
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 652 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-559-896B-2

Query Match          4.1% Score 98; DB 4; Length 652;
Best Local Similarity 21.8%; Pred. No. 0.15;
Matches 41; Conservative 34; Mismatches 71; Indels 42; Gaps 7;

QY 265 OKROFYSHRQPMALQVMSDRSDEVDVADPEDROMLDPVYVNDK 315
DB 289 EEKEKIHREKIKERKINKMDQIDKITEELINKMSDEICHVRAILE--DIQKE 345
QY 316 EKQPMILMSFYRKQRYADGHISWACEAFSFEKEKELHRYSSLEFCWRLFLIKLNNHGL 375
DB 346 KIQNLEL-----EIDRLKYEELDRDR---EAREIIPMRN-- 378
QY 376 VDSATINNCNTILENCNSSDTTNNNSVDRPSDSTNNNTVDHPNDINKNNVNDK 435
DB 379 --LSRNEKNDIIRKINKESNOKRKEVNVFIIDNNDNSNNNNNNRRVNNLNT--KHT 435
QY 436 DNNSRDV 443
DB 436 NNNYENV 443

RESULT 12
US-08-446-855A-2
; Sequence 2, Application US/08446855A
; Patent No. 5849573
; GENERAL INFORMATION:
; APPLICANT: Stewart, Thomas S
; APPLICANT: Flores, Maria V
; APPLICANT: O'Sullivan, William J
; TITLE OF INVENTION: Nucleotide sequence encoding carbamoyl
; EARLIER FILING DATE: 1995-07-06
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
```

```
; STREET: 1100 No. 5849573th Gleebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent In Release #1.24
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/446,855A
; FILING DATE: 06-Jul-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Mitchard, Leonard C
; REGISTRATION NUMBER: 29,009
; REFERENCE/DOCKET NUMBER: 47-80
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2391 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-446-855A-2

Query Match          4.0% Score 94; DB 2; Length 2391;
Best Local Similarity 21.4%; Pred. No. 3.4;
Matches 44; Conservative 40; Mismatches 64; Indels 58; Gaps 12;

QY 266 KROFYSHRQPMALQVMSDRSDEVDVADPEDROMLDPVYVNDKQFQFHL 322
DB 269 KEFNYIN-----EITNDSSMEDHDNINISINFNCCSISF-DKSKRYVINH- 319
QY 323 WNSFYR-KQRYIAGHISWACEAFSFEKEELH-----RYSLEFCWRLFLIKLNN 372
DB 320 --TLIKDKMLITS-----SEETLIDLNCFSSSDKDSF--KLIGICEYD 364
QY 373 HGLVD-----SATINNCN-----TILENCNSSDTTNNNSVDRPSDSTNNNT-- 419
DB 365 KYLIDLEBNASFHYNNVDEYGYDVNKNINILSNKIEQNNNNNNNNNNNEVDYI 424
QY 420 -VDHPNDINK-----NNVDNKN 437
DB 425 KKDEBNVNSKVFSQYNNNAONNEH 450

RESULT 13
US-09-150-741-2
; Sequence 2, Application US/09150741
; Patent No. 6183996
; GENERAL INFORMATION:
; APPLICANT: Stewart et al.
; TITLE OF INVENTION: Nucleotide Sequence Encoding Carbamoyl Phosphate
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/150,741
; EARLIER FILING DATE: 1998-09-10
; EARLIER APPLICATION NUMBER: PI6380
; EARLIER FILING DATE: 1997-12-16
; EARLIER APPLICATION NUMBER: A093/00617
; EARLIER FILING DATE: 1995-12-02
; EARLIER APPLICATION NUMBER: 08/446,855
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2391
```

```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: protein
US-09-150-741-2

```

Query Match 4.0%; Score 94; DB 4; Length 2391;

Best Local Similarity 21.4%; Pred. No. 3.4; Matches 44; Conservative 40; Mismatches 64; Indels 58; Gaps 12;

```

QY 266 KRQYHSRVPQMALEQVNSDRSD---EYDDVADFEQMLDEVYVNDKQFMHL 322
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 269 KEEFYTN-----EMITNDSMEDHDNEINGSISNFNCPSISF--DKSEKVIYH- 319
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 323 WMSFYR-KQVIADGHISACAFSRYKEH-----RYSLEFCMRLEFLIKLN 372
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 320 -TLRDKMNLTS-----SEELYKDLHNCNFSNSDKNSFF--KLYGICEYD 364
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 373 HGLVD-----SATINNCN-----TILENCRNSSDTTNNNSVDRPSDNTNNNI- 419
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 365 KLIIDLEENASPHYNNVDEYGYDVNKNNTNIIISNKKIEQNNNNNNKNNNNNEVDYI 424
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
QY 420 -VDHPNDINNK-----NNVDNKN 437
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 425 KDEDNNVNSKVYFSQYNNNAQNNH 450
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |

```

RESULT 14

```

US-09-384-162-6
; Sequence 6, Application US/09384162
; Patent No. 6376747
; GENERAL INFORMATION:
; APPLICANT: Xing, Ti
; APPLICANT: Malik, Kamal
; APPLICANT: Martin-Heller, Teresa
; APPLICANT: Miki L., Brian
; TITLE OF INVENTION: No. 6376747el Plant-Derived Map Kinase Kinase
; FILE REFERENCE: 08-884280US
; CURRENT APPLICATION NUMBER: US/09/384,162
; CURRENT FILING DATE: 1999-08-27
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 286
; TYPE: PRT
; ORGANISM: Dictyostellium discoideum
US-09-384-162-6

```

Query Match

3.9%; Score 92; DB 4; Length 286;

Best Local Similarity 37.0%; Pred. No. 0.18; Matches 20; Conservative 10; Mismatches 14; Indels 10; Gaps 2;

```

QY 381 INNCTILENCNRSDDTTTNNNSVDRPSDNTNNNIIVDHPNDINNKNNVDN 434
   | | | : : : | : : : | : : : | : : : | : : : | : : : |
Db 195 INNSNNINIRSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNVNL-----DISNGGLVDS 238

```

RESULT 15

```

US-09-177-249-2
; Sequence 2, Application US/09177249
; Patent No. 6229064
; GENERAL INFORMATION:
; APPLICANT: Fischer, Robert L.
; APPLICANT: Chad, Nir
; APPLICANT: Kiyosue, Tomohiro
; APPLICANT: Yadegari, Ramih
; APPLICANT: Margossian, Linda
; APPLICANT: Harada, John
; APPLICANT: Goldberg, Robert B.
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids That Control Seed and Fruit
; FILE REFERENCE: 023070-086120US

```

```

; CURRENT APPLICATION NUMBER: US/09/177,249
; CURRENT FILING DATE: 1998-10-22
; EARLIER APPLICATION NUMBER: US 09/071,838
; EARLIER FILING DATE: 1998-05-01
; NUMBER OF SEQ ID NOS: 324
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Arabidopsis sp.
US-09-177-249-2

```

Query Match 3.8%; Score 91.5; DB 4; Length 689;

Best Local Similarity 20.2%; Pred. No. 0.84; Matches 49; Conservative 39; Mismatches 61; Indels 93; Gaps 11;

```

QY 246 TTRKLSARSEARSHLLQKQFYHSHVQPMALQVNSDRSDEVDVADFEQFM 305
   | : : | : : | : : | : : | : : | : : | : : | : : |
Db 142 TKSQLMAL-----SDSVIGKQIYY--LNGEAL-LSSEDEDEDEDE- 183
   | : : | : : | : : | : : | : : | : : | : : | : : |
QY 306 LDDFYDVNDKQFQFHLNNSFYRKQVIADGHISAC-----EAFSRYEKE- 352
   | : : | : : | : : | : : | : : | : : | : : | : : |
Db 184 ---EIKKEKCEFSQVDFI-----WVGQDYGIDLVVRALAKYLEVDV 227
   | : : | : : | : : | : : | : : | : : | : : | : : |
QY 353 ---LHRYSLFCQWRLEFLIKLNNHGLVDSATINNCNTIL-----ENCN- 393
   | : : | : : | : : | : : | : : | : : | : : | : : |
Db 228 SILERYNEI-----KLKNDGTAGASDLTSTTTTAFQDFADRRHRCRMFDC 278
   | : : | : : | : : | : : | : : | : : | : : | : : |
QY 394 -----SSDTTNNNSVDRPSD-----SNTNNNIYDHPNDINNKNNVDN 435
   | : : | : : | : : | : : | : : | : : | : : | : : |
Db 279 MHEKYEPESSSEDSKSLFDEDDROPSCSEHCYLVKRSYTEADHYMDNDNSISNKIYVSDP 338
   | : : | : : | : : | : : | : : | : : | : : | : : |
QY 436 DN 437
   | : : | : : | : : | : : | : : | : : | : : | : : |
Db 339 NN 340

```

Search completed: June 13, 2003, 15:45:55
Job time : 28 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:34:30 ; Search time 25 Seconds
(without alignments)
1903.099 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRQNCRAKSSPEEYISDPE.....INKNKNVDKNQNSDKYIK 445

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 408643 seqs, 106915682 residues

Total number of hits satisfying chosen parameters: 408643

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCF_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/PCFUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	225	9.5	388	10 US-09-764-864-931	Sequence 931, App
2	221	9.3	739	9 US-09-874-162A-5	Sequence 5, Appl
3	221	9.3	776	9 US-09-874-162A-8	Sequence 8, Appl
4	153	6.4	289	10 US-09-764-864-911	Sequence 911, App
5	127.5	5.4	961	10 US-09-801-368-132	Sequence 132, App
6	126.5	5.3	2150	9 US-10-135-322-17	Sequence 17, Appl
7	125.5	5.3	1501	10 US-09-924-154-17	Sequence 17, Appl
8	124	5.2	1288	9 US-10-087-780-2	Sequence 2, Appl
9	122.5	5.2	1331	9 US-10-087-780-50	Sequence 50, Appl
10	115	4.8	175	10 US-09-764-864-1355	Sequence 1355, Ap
11	113	4.8	807	9 US-09-820-843A-108	Sequence 108, App
12	111.5	4.7	1143	10 US-09-924-154-14	Sequence 14, Appl
13	105	4.4	732	9 US-09-832-292-12	Sequence 12, Appl
14	105	4.4	732	10 US-09-994-485-8	Sequence 8, Appl
15	104.5	4.4	97	9 US-09-858-935B-77	Sequence 77, Appl
16	104.5	4.4	666	10 US-09-801-368-36	Sequence 36, Appl
17	101	4.2	534	10 US-09-801-368-124	Sequence 124, App
18	100.5	4.2	97	9 US-09-858-935B-72	Sequence 72, Appl
19	99.5	4.2	292	10 US-09-764-864-1369	Sequence 1369, Ap

20	99.5	4.2	861	9 US-09-820-843A-109	Sequence 109, App
21	98	4.1	97	9 US-09-858-935B-76	Sequence 76, Appl
22	98	4.1	652	10 US-09-351-794A-2	Sequence 2, Appl
23	97.5	4.1	97	9 US-09-858-935B-73	Sequence 73, Appl
24	97.5	4.1	97	9 US-09-858-935B-74	Sequence 74, Appl
25	97.5	4.1	97	9 US-09-858-935B-75	Sequence 75, Appl
26	97	4.1	440	9 US-10-083-357-1266	Sequence 1266, Ap
27	97	4.1	559	10 US-09-801-368-256	Sequence 256, App
28	96.5	4.1	97	9 US-09-858-935B-70	Sequence 70, Appl
29	96.5	4.1	97	9 US-09-858-935B-71	Sequence 71, Appl
30	95	4.0	1093	10 US-09-801-368-392	Sequence 392, App
31	94.5	4.0	393	9 US-10-153-668-76	Sequence 76, Appl
32	94.5	4.0	393	9 US-10-153-668-78	Sequence 78, Appl
33	94.5	4.0	556	9 US-10-012-896-1005	Sequence 1005, Ap
34	94.5	4.0	1610	9 US-10-155-533-9	Sequence 9, Appl
35	93.5	3.9	9	9 US-09-820-843A-32	Sequence 32, Appl
36	93	3.9	676	10 US-09-801-368-302	Sequence 302, App
37	93	3.9	2086	10 US-09-815-242-5639	Sequence 5639, Ap
38	93	3.9	5795	10 US-09-815-242-12610	Sequence 12610, A
39	92.5	3.9	704	10 US-09-801-368-218	Sequence 218, App
40	92.5	3.9	758	10 US-09-801-368-224	Sequence 224, App
41	92.5	3.9	762	9 US-09-738-628-4825	Sequence 4825, Ap
42	91.5	3.8	91	9 US-09-858-935B-69	Sequence 69, Appl
43	91.5	3.8	689	10 US-09-071-838-2	Sequence 2, Appl
44	91.5	3.8	1786	9 US-09-742-096-3	Sequence 3, Appl
45	91	3.8	1051	10 US-09-911-888-14	Sequence 14, Appl

ALIGNMENTS

RESULT 1		US-09-764-864-931	
Sequence 931, Application US/09764864		Patent No. US20020132753A1	
GENERAL INFORMATION:		APPLICANT: Rosen et al.	
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies		FILE REFERENCE: PT223	
CURRENT APPLICATION NUMBER: US/09/764,864		CURRENT FILING DATE: 2001-01-17	
Prior application data removed - consult PALM or file wrapper		NUMBER OF SEQ ID NOS: 1792	
SOFTWARE: PatentIn Ver. 2.0		SEQ ID NO 931	
LENGTH: 388		TYPE: PRT	
ORGANISM: Homo sapiens		US-09-764-864-931	
Query Match		9.5%: Score 225; DB 10; Length 388;	
Best Local Similarity 21.1%: Pred. No. 6, 2e-13;		Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;	
QY	58 RRRK-----SRSTGVVFNKDCNNTLTQTEFEDSCSPFCSMILGSEFGKLOFHLNSSHD	112	Sequence 931, App
DB	62 RREKTPNNRNKRLNIFVFLYNNMTROTETARDLHCWCTLNCRLKLSLKLKILCHS	121	Sequence 8, Appl
QY	113 LFEFEFKLFEEYQTVNVSVKLSFTEEGSDDDKFEPPSLCSKPRK-RRRG--GRNN	168	Sequence 132, App
DB	122 RIFVFVYHPKARDIVSI-----NKGYG-----SYAGNVODIHRRGFAPSRNG	167	Sequence 17, Appl
QY	169 -TRRLKVCPLPDSPLFNGTENGITLLDNGRGIGYPATLADGFEMTSNIPALNHS	227	Sequence 2, Appl
DB	168 PVKRPPIHILVCRRKRTKASSEPLESDG-----	198	Sequence 50, Appl
QY	228 SLDAKAVLTSEAVPATKTRKLSAERSEARSHLLQKRFYHSHRVPALAEQVMSDR	287	Sequence 108, App
DB	199 -----EVEQQRYSNGH-----NRLTFHNDJCLPLRQDM--EV	230	Sequence 14, Appl
QY	288 DSEDEVDDVADPEDROMIDFVDVKNCKEQPMHLNVSFRKORYIADGHSWACEAFSR	347	Sequence 77, Appl
DB	231 DSEKDEPEMLREKTTITQTEFFSDVNGEKEVYMKLMLNLMVHKHGFIALDNQNMHACMLFE	290	Sequence 36, Appl

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rapb

Page 2

QY 348 FYEKELHRYSSLFMCWRLFLIKLNHGLVDATINNCNTILLENCRSSDPTTTNNNSVD 407
DB 291 NYGOKLIK-KNLCRNFMHLVSMHDFNLISIMSIDKAVTKLRMO-----OKLE 338
QY 408 RPSDSNTNNNIYDHPN-----DINNKNVNDKDNNS 439
DB 339 KGESASPAHEITEEONGTANGSEINSKEKALETDSVS 377

RESULT 2
US-09-874-162A-5
; Sequence 5, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: Koonitz, Jason
; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAZF1 AND JJA1 GENES IN
; FILE REFERENCE: 05311-024001
; CURRENT APPLICATION NUMBER: US/09/874,162A
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,093
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 739
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-874-162A-5

Query Match 9.3%; Score 221; DB 9; Length 739;

Best Local Similarity 21.1%; Pred. No. 3.8e-12;

Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

QY 58 RRRK-----SRSTGVVFNKYDCNNLTOKTEVEDSCSPFCMSLGSFGLOPHLNSHD 112
DB 413 RREKDPNNRQKRLIFQFLYNNNTROQTEARDLHCPCWCTLNCRLKLSLKHLKCHS 472
QY 113 LFEFEFKLFEEYQTVNVSVKLSNFIFFEEGSDDDKFEFSLCSKPRK--RRRG--GRNN 168
DB 473 RIFENVYVHPKGRIDVSI-----NECYDG-----SYAGNPQDIHRQGFAPFSRNG 518
QY 169 -TRRLKVCFLPDSPLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHS 227
DB 519 PVKRTPIITHILVCRPKRTKASMSFELESDG----- 549
QY 228 SLIDAGAKVILTSEAVVPATRTKRLSAERSEARSHLLQKRFYHSHRVQPMALBOVMSDR 287
DB 550 -----EVEDORTYSSGH-----NRLYFHSPTCLPLRPOEM--EV 581
QY 288 DSEDEVDVADFEEDROMLDDFVYVNDKQFHLNNSFYKORVIADGHISWACEAFSR 347
DB 562 DSEDEKDEPMLREKTIITQIEFSDVNEGEKVKMLNLMHMGFTADQNMHACMLFEVE 641
QY 348 FYEKELHRYSSLFMCWRLFLIKLNHGLVDATINNCNTILLENCRSSDPTTTNNNSVD 407
DB 642 NYGOKLIK-KNLCRNFMHLVSMHDFNLISIMSIDKAVTKLRMO-----OKLE 689
QY 408 RPSDSNTNNNIYDHPN-----DINNKNVNDKDNNS 439
DB 690 KGESASPAHEITEEONGTANGSEINSKEKALETDSVS 728

RESULT 3

US-09-874-162A-8

; Sequence 8, Application US/09874162A
; Patent No. US20020155452A1

; GENERAL INFORMATION:
; APPLICANT: Koonitz, Jason

; APPLICANT: Sklar, Jeffrey
; TITLE OF INVENTION: FUSION OF JAZF1 AND JJA1 GENES IN

; TITLE OF INVENTION: ENDOMETRIAL STROMAL TUMORS
; FILE REFERENCE: 05311-024001
; CURRENT APPLICATION NUMBER: US/09/874,162A
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,093
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-874-162A-8

Query Match 9.3%; Score 221; DB 9; Length 776;

Best Local Similarity 21.1%; Pred. No. 4.1e-12;

Matches 84; Conservative 64; Mismatches 151; Indels 100; Gaps 12;

QY 58 RRRK-----SRSTGVVFNKYDCNNLTOKTEVEDSCSPFCMSLGSFGLOPHLNSHD 112
DB 450 RREKDPNNRQKRLIFQFLYNNNTROQTEARDLHCPCWCTLNCRLKLSLKHLKCHS 509
QY 113 LFEFEFKLFEEYQTVNVSVKLSNFIFFEEGSDDDKFEFSLCSKPRK--RRRG--GRNN 168
DB 510 RIFENVYVHPKGRIDVSI-----NECYDG-----SYAGNPQDIHRQGFAPFSRNG 555
QY 169 -TRRLKVCFLPDSPLTNGTENGITLLDNGNGLGYPEATELAGOFEMTSNIPPAIHS 227
DB 556 PVKRTPIITHILVCRPKRTKASMSFELESDG----- 586
QY 228 SLIDAGAKVILTSEAVVPATRTKRLSAERSEARSHLLQKRFYHSHRVQPMALBOVMSDR 287
DB 587 -----EVEDORTYSSGH-----NRLYFHSPTCLPLRPOEM--EV 618
QY 288 DSEDEVDVADFEEDROMLDDFVYVNDKQFHLNNSFYKORVIADGHISWACEAFSR 347
DB 619 DSEDEKDEPMLREKTIITQIEFSDVNEGEKVKMLNLMHMGFTADQNMHACMLFEVE 678
QY 348 FYEKELHRYSSLFMCWRLFLIKLNHGLVDATINNCNTILLENCRSSDPTTTNNNSVD 407
DB 679 NYGOKLIK-KNLCRNFMHLVSMHDFNLISIMSIDKAVTKLRMO-----OKLE 726
QY 408 RPSDSNTNNNIYDHPN-----DINNKNVNDKDNNS 439
DB 727 KGESASPAHEITEEONGTANGSEINSKEKALETDSVS 765

RESULT 4

US-09-764-864-911

; Sequence 911, Application US/09764864
; Patent No. US2002013753A1

; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT223

; CURRENT APPLICATION NUMBER: US/09/764,864
; PRIOR FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 911

; LENGTH: 289
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-864-911

Query Match 6.4%; Score 153; DB 10; Length 289;

Best Local Similarity 23.9%; Pred. No. 2.7e-06;

Matches 49; Conservative 43; Mismatches 89; Indels 24; Gaps 5;

QY 245 ATKTRKLSAENSEA---KSHLLQKRFYHSHRVQPMALBOVMSDSDSEVDVADFE 301
DB 88 ASMSFELESDGEVHOQRTYSSGHNRLYFHSPTCLPLRPOEM--EVDSEKDEPMLRREK 145

QY 302 DROMLDFVYNKDEKOFHMLNMFVRKQKVADGHSNACFAFSFYKELHRYSLEW 361
Db 146 TITQIEFSDVNEGEVEVKMLNMLVMAKGFADNOMACMLFVNTYGKLIK-KNLCR 204
QY 362 CWRLFLIKMLNHLVDATINNCNTLNCNRSDDTTNNNSVDRPSDSTNNNIYD 421
Db 205 NFPLHLVSMHDFLJISIMSDKAVTKLREMQ-----QKLEKGSASAPANEITE 253
QY 422 HRP-----DINNKNNVDKNNS 439
Db 254 EQNGTANGFSEINSKEKALETDSVS 278

RESULT 5
US-09-801-368-132
; Sequence 132, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132
; LENGTH: 961
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-132

Query Match 5.4%; Score 127.5; DB 10; Length 961;
Best Local Similarity 41.3%; Pred. No. 0.0041;

Matches 26; Conservative 13; Mismatches 23; Indels 1; Gaps 1;

QY 382 NNCNTLNCNRSDDTTNNNSVDRPSDSTNNNIYDHPDINN-KNVDKNNSR 440
Db 501 NNNNNNDNDNNNSNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 560
QY 441 DKY 443
Db 561 DNT 563

RESULT 6
US-10-135-322-17
; Sequence 17, Application US/10135322
; Patent No. US2002017301A1
; GENERAL INFORMATION:
; APPLICANT: BENEFY, PN
; APPLICANT: HELARIUTTA, Y
; APPLICANT: MAHONEN, AP
; APPLICANT: BONKE, AMM
; APPLICANT: KAUPPINEN, L
; APPLICANT: RIKONEN, M
; TITLE OF INVENTION: WOODEN LEG GENE, PROMOTER AND USES THEREOF

; FILE REFERENCE: 5914-086-999
; CURRENT APPLICATION NUMBER: US/10/135,322
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: 60/253,739
; PRIOR FILING DATE: 2000-11-29
; NUMBER OF SEQ ID NOS: 43
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 17
; LENGTH: 2150
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-10-135-322-17

Query Match 5.3%; Score 126.5; DB 9; Length 2150;
Best Local Similarity 17.5%; Pred. No. 0.017; Indels 163; Gaps 16;
Matches 85; Conservative 64; Mismatches 175;

QY 2 CROCRKAS-----SPEEVISTDENLTIYCKPYALYNIFHLRSIGN-----PSFLPCLNY 52
Db 353 CGSNGSGNGGIPPLSPRLNLSLNGVAVSPRNI-----HLNNLNNSNMLPLSPRLHNF 406
QY 53 KIGAKRRKRSRSTGAVYVYKDCNNTLOKTEVREDCSCFPCSMLCGFFGLQPHLSSHD 112
Db 407 HINVSMLNN 435
QY 113 LFEFEFKLFEYQTVAVSVKLSNFIPEEGSDDDKFEPPSLCKPRKRRGRG 166
Db 436 -----NVSRRNNHNSIPRGSN--ISP-----RSNNGGSTITSPR 468
QY 167 -----NTRRLKVCFLPLDSPSL--TNGTENGITLLDNGRGLGYPEA-TELADGP 214
Db 469 NISNNNNIINNINNMLILPPRNSPRLNVNPNISRLATLSNLTPLVSSLTSSNNNN 528
QY 215 EMTSNIPALIASS-----LDGAKVILITSEAVV--PATKTRKLSAESEARSHLLQK 266
Db 529 QSNNTNPSPINNNGNGHICQIPISEILGNKYVYVNNNNNNNNNNNNNNNNNNNN 588
QY 267 RQYHSRYPQMALEQVMSROSEDEVDVADFEORML--DDFYD-----VVK 314
Db 589 NNN 648
QY 315 DEKOFHMLNMFVRKQKVADGHSNACFAFSFYKELHRYSLEWCMRLFLIKMLNG 374
Db 649 DDND-----ENDG 656
QY 375 LVDSATINNCNTLNCNRSDDTTNNNSVDRPSDSTNNNIYDHPDINNKNVDN 434
Db 657 NSNNTSNNSNN 716
QY 435 KDNNSRD 441
Db 717 NNNNNNN 723

RESULT 7
US-09-924-154-17
; Sequence 17, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Natum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17
; LENGTH: 1501
; TYPE: PRT
; ORGANISM: Mammalian


```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (4)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (136)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (166)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1355
```

Query Match 4.8%; Score 115; DB 10; Length 175;

Best Local Similarity 20.4%; Pred. No. 0.0053; Matches 46; Conservative 30; Mismatches 74; Indels 76; Gaps 8;

```

QY 76 NNTLQKTEVREDSCSPFCSMICGSPKGLQPHLNSHDLFEFEFKLFEYQYVNVSKLNS 135
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 8 NNTROOTEARDLHCPWCTLNCRLKLSLKHLKLCRSRTFNVVHPKCARIDVSI----- 63
QY 136 FTFEEBGSDDKFEFPFSLCSKPRK-RRQRC--GRNN-TRRLKVCFLPLDSPSLNGTEN 190
   64 ---NECYDG-----SYAGNPQDIHRQGFASRNGPYKRPITHILICRPRKTKASMS 113
QY 191 GITLNDGNGRLGYPEATELAGOFEMTSNIPALIAHSSLDAGAKVILTSEAVVPATKTKR 250
   ||
DB 114 ELESEDEG-----E 122
QY 251 ISAESEARSHLLQKROFYHSHRYQPMALQVMSDRSEDEVD 296
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB 123 VEQQRYSQGH---NRXYFHSDFCLPLRQEM--EVDSEDEKPE 162
```

RESULT 11

```

US-09-820-843A-108
; Sequence 108, Application US/09820843A
; Publication No. US20030039963A1
; GENERAL INFORMATION:
; APPLICANT: Council of Scientific and Industrial Research
; TITLE OF INVENTION: A COMPUTATIONAL METHOD FOR THE IDENTIFICATION OF CANDIDATE PROTEIN
; FILE REFERENCE: 063915
; CURRENT APPLICATION NUMBER: US/09/820,843A
; CURRENT FILING DATE: 2001-03-30
; NUMBER OF SEQ ID NOS: 118
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 108
; LENGTH: 807
; TYPE: PRT
; ORGANISM: Plasmodium falciparum
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: hypothetical protein
; NAME/KEY: misc_feature
; OTHER INFORMATION: g113845292
US-09-820-843A-108
```

Query Match 4.8%; Score 113; DB 9; Length 807;
Best Local Similarity 35.3%; Pred. No. 0.076;
Matches 24; Conservative 14; Mismatches 24; Indels 6; Gaps 1;

```

QY 372 NNGIVDSATINNCNTILENCRSSDPTTNNNNNSVDSPDSNTNNNIYDHPDINKNN 431
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 157 NYGLKRRKITLRLDKGYNNEITLNNKNNLNKNNNYDNRN-----NNNNKNN 210
QY 432 VDNKDNN 439
   ::|||::
DB 211 INNNNNN 218
```

RESULT 12

```

US-09-924-154-14
; Sequence 14, Application US/09924154
; Patent No. US20020127241A1
; GENERAL INFORMATION:
; APPLICANT: Natrum, David L.
; APPLICANT: Sim, Kim L.
; TITLE OF INVENTION: Anti-Plasmodium Compositions and Methods of Use
; FILE REFERENCE: 05213-0465 43170-262105
; CURRENT APPLICATION NUMBER: US/09/924,154
; CURRENT FILING DATE: 2001-08-07
; PRIOR APPLICATION NUMBER: US 60/223,525
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 1143
; TYPE: PRT
; ORGANISM: Mammalian
US-09-924-154-14
```

Query Match 4.7%; Score 111.5; DB 10; Length 1143;

Best Local Similarity 18.6%; Pred. No. 0.17; Matches 94; Conservative 70; Mismatches 171; Indels 171; Gaps 24;

```

QY 2 CRONCR-----AKSPSEYI--STDENLLTYCKPVRLYNIPIHLSIGNPSFLPRL 50
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 664 CNKCKSEYKMWIDLAKKSEYKQVDKTKDNKKMY-----DNIDVK----- 705
QY 51 NTKIGAKRRKRSRSTGMVFNKDCNNTLQKTEVREDSCSPFCMKGSPKGLQPHLN-- 108
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB 706 NKEANVYLEKSKCKCDVNFDDKIFNES--PNEYEDMK-----KCEIK--YLNBI 753
QY 109 ---SSHDLFEFEFKLFEYQYVNVSVKNSFIFEBSGDD-----KFEPE 151
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 754 KYPKTKHDITDID--TFSTPFGDGPISINANINEDQSGKDSNTGNSPVSHEPE 811
QY 152 S-----LCSKPRKRQRCGRNNTRLKVCFLPLDSPSLNGTENGITLNDGNGRLG 203
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 812 SDAINVEKLSGDSESSSTRG-----ILDINDPSVNNVNE----- 847
QY 204 YPATELAGOFEMTSNIPALIAHSSLD--AGAKVILTSEAVVPATKTKLSERARSH 261
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB 848 VHDASNTQGSVNTSDITNGHSESSLNRTTNQDIDIKRSGNEQSDNQSSSHSDNSGS 907
QY 262 LLL-----QKROFYHS--HRVOPMALQVMSD-----RSEDEVDVDD 299
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 908 LITIGVPSBDNTONTYDSQNPFRDTPNALASLPSDDKINEIEGFDSSRSGEGRGDTTSN 967
QY 300 FEDRQMLDPEVDVNVKDEKQFMHLMNSFVRKORYIADGHISWACEAFSFEYERELHRYSSL 359
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 968 THDVR-----TNVSRRRVNSH--DFINGMANNAH-----HOY----- 1001
QY 360 FWCWRLFILKLMNHGLVDSATINNCNTILENCRSSDPTTNNNNNSVDSPDSNTNNNI 419
   :|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::
DB 1002 -----TVOIENNGIIT-----RCGEESAGSNVYKDKMPKSNFSSEN- 1037
QY 420 VDHDPDINKNNVNDKNNSRDKYK 445
   |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::
DB 1038 -DHKNIOEYNSRDKY--RVREIIR 1060
```

RESULT 13

```

US-09-832-292-12
; Sequence 12, Application US/09832292
; Patent No. US20020177205A1
; GENERAL INFORMATION:
; APPLICANT: Ryzanov, Alexey
; TITLE OF INVENTION: MAMMALIAN ALPHA-KINASE PROTEINS, NUCLEIC ACIDS AND
; FILE REFERENCE: 601-1-098CIP
; CURRENT APPLICATION NUMBER: US/09/832,292
; CURRENT FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 09/632,131
```

Fri Jun 20 09:03:39 2003

us-09-890-220-2.rapb

Page 6

PRIOR FILING DATE: 2001-08-03
NUMBER OF SEQ ID NOS: 45
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 732
TYPE: PRT
ORGANISM: Dictyostellium discoideum
US-09-832-292-12

Query Match
Best Local Similarity 20.1%, Pred. No. 0.38;
Matches 76; Conservative 61; Mismatches 167; Indels 74; Gaps 17;

QY 117 EFKLFE-EYQTVNVSVKLNSFIFFEEG---SDDKFEFSLCSKPKRRORGRNNTRL 172
| | | | | : : : : : | | | | : : : : :
Db 34 EGIKIFRPGYGTSDLDIGKMLPGSGGATADSKFEKFAKNTLADIQKGDLLYAV 93
| | | | | : : : : : | | | | : : : : :
QY 173 KV-----CFLPDSPLTNGTENGITLLDNGRGLGYPATELAGOFEWTSNI-----PP 222
| | | | | : : : : : | | | | : : : : :
Db 94 KSKRPTNDSLPTLINIAFLDGSERAIKWEYD-----PYTT--TAQWCTATLYKVEPV 144
| | | | | : : : : : | | | | : : : : :
QY 223 AIAHSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLQKRF-----YHSIR 274
| | | | | : : : : : | | | | : : : : :
Db 145 PFAEGAFKRAHTLDLSKSGASGRVSKIGKTPRPSYEDVKMOMAKKADKINSFK 204
| | | | | : : : : : | | | | : : : : :
QY 275 VOPMALEQVMS-----DRSEDEV--DDVADFEEDQMLDDEVYDNKDEK---QFMHL 322
| | | | | : : : : : | | | | : : : : :
Db 205 -PPKIEFLQSCVLEFVDRTSDDLICGAEPYEGQYRKYNNSGFVSNDEKRTPOSFSHF 263
| | | | | : : : : : | | | | : : : : :
QY 323 WNSFVKQKVIAD-----GHISWACEAFSFEYKELHRSLSLEFCWRLFLIKLNMHGLVDS 378
| | | | | : : : : : | | | | : : : : :
Db 264 TYESHNLHLLIDIGVG-----DHYTDPIHTYDGV---GFGIGNIGQKGEFEK 310
| | | | | : : : : : | | | | : : : : :
QY 379 ATINNCNTILE--NCRNSDPTTTNNNSVDPRS-----DSNTNNNTVDHPNDIN 427
| | | | | : : : : : | | | | : : : : :
Db 311 LDTHKCMALCOYLNLQSLNPKSEKSDGTVPRDLIFPDTSEBDNNNNNN--NNNNNN 368
| | | | | : : : : : | | | | : : : : :
QY 428 NKNVNDKDNNSRDKVIR 445
| | | | | : : : : : | | | | : : : : :
Db 369 NNNNNNNNNNNSSISK 386
| | | | | : : : : : | | | | : : : : :

RESULT 14
US-09-994-485-8
Sequence 8, Application US/09994485
Patent No. US20020142429A1
GENERAL INFORMATION:
APPLICANT: Ryazanov, Alexey G.
Halt, William N.
Pavur, Karen S.
TITLE OF INVENTION: ELONGATION FACTOR-2 KINASE (EF-2 KINASE)
AND METHODS OF USE THEREFOR
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: David A. Jackson, Esq.
STREET: 411 Hackensack Ave, Continental Plaza, 4th
FLOOR
CITY: Hackensack
STATE: New Jersey
COUNTRY: USA
ZIP: 07601
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/09/994,485
APPLICATION NUMBER: US/09/994,485
FILING DATE: 27-NO. US20020142429A1-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/914,999
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Jackson Esq., David A.
REGISTRATION NUMBER: 26,742
REFERENCE/DOCKET NUMBER: 601-1-078
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-487-5800
TELEFAX: 201-343-1684
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 732 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: NO
ORIGINAL SOURCE:
ORGANISM: Dictyostellium discoideum
SEQUENCE DESCRIPTION: SEQ ID NO: 8:
US-09-994-485-8

Query Match
Best Local Similarity 4.4%, Score 105; DB 10; Length 732;
Matches 76; Conservative 61; Mismatches 167; Indels 74; Gaps 17;

QY 117 EFKLFE-EYQTVNVSVKLNSFIFFEEG---SDDKFEFSLCSKPKRRORGRNNTRL 172
| | | | | : : : : : | | | | : : : : :
Db 34 EGIKIFRPGYGTSDLDIGKMLPGSGGATADSKFEKFAKNTLADIQKGDLLYAV 93
| | | | | : : : : : | | | | : : : : :
QY 173 KV-----CFLPDSPLTNGTENGITLLDNGRGLGYPATELAGOFEWTSNI-----PP 222
| | | | | : : : : : | | | | : : : : :
Db 94 KSKRPTNDSLPTLINIAFLDGSERAIKWEYD-----PYTT--TAQWCTATLYKVEPV 144
| | | | | : : : : : | | | | : : : : :
QY 223 AIAHSLDAGAKVILTSEAVVPATKTRKLSAERSEARSHLLQKRF-----YHSIR 274
| | | | | : : : : : | | | | : : : : :
Db 145 PFAEGAFKRAHTLDLSKSGASGRVSKIGKTPRPSYEDVKMOMAKKADKINSFK 204
| | | | | : : : : : | | | | : : : : :
QY 275 VOPMALEQVMS-----DRSEDEV--DDVADFEEDQMLDDEVYDNKDEK---QFMHL 322
| | | | | : : : : : | | | | : : : : :
Db 205 -PPKIEFLQSCVLEFVDRTSDDLICGAEPYEGQYRKYNNSGFVSNDEKRTPOSFSHF 263
| | | | | : : : : : | | | | : : : : :
QY 323 WNSFVKQKVIAD-----GHISWACEAFSFEYKELHRSLSLEFCWRLFLIKLNMHGLVDS 378
| | | | | : : : : : | | | | : : : : :
Db 264 TYESHNLHLLIDIGVG-----DHYTDPIHTYDGV---GFGIGNIGQKGEFEK 310
| | | | | : : : : : | | | | : : : : :
QY 379 ATINNCNTILE--NCRNSDPTTTNNNSVDPRS-----DSNTNNNTVDHPNDIN 427
| | | | | : : : : : | | | | : : : : :
Db 311 LDTHKCMALCOYLNLQSLNPKSEKSDGTVPRDLIFPDTSEBDNNNNNN--NNNNNN 368
| | | | | : : : : : | | | | : : : : :
QY 428 NKNVNDKDNNSRDKVIR 445
| | | | | : : : : : | | | | : : : : :
Db 369 NNNNNNNNNNNSSISK 386
| | | | | : : : : : | | | | : : : : :

RESULT 15
US-09-858-935B-77
Sequence 77, Application US/09858935B
Publication No. US2003006917A1
GENERAL INFORMATION:
APPLICANT: Dubaquitte, Yves
APPLICANT: Lowman, Henry B.
TITLE OF INVENTION: METHOD FOR TREATING CARTILAGE DISORDERS
FILE REFERENCE: P1794R1
CURRENT APPLICATION NUMBER: US/09/858,935B
APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: US 60/248,985
PRIOR FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 60/204,490
PRIOR FILING DATE: 2000-05-16
NUMBER OF SEQ ID NOS: 153
SEQ ID NO 77
LENGTH: 97
TYPE: PRT

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:35:49 ; Search time 310 Seconds

(without alignments)
925.504 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRKSSPEEVISTDENL...INNNKNDKNDKNDKDKYK 445

Scoring table:

BLOSUM62

Searched: 4569144 seqs, 64473110 residues

Total number of hits satisfying chosen parameters: 4569144

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCITUS.COMB.pep:*
2: /cgn2_6/ptodata/1/paa/US06.COMB.pep:*
3: /cgn2_6/ptodata/1/paa/US07.COMB.pep:*
4: /cgn2_6/ptodata/1/paa/US08.COMB.pep:*
5: /cgn2_6/ptodata/1/paa/US081.COMB.pep:*
6: /cgn2_6/ptodata/1/paa/US082.COMB.pep:*
7: /cgn2_6/ptodata/1/paa/US083.COMB.pep:*
8: /cgn2_6/ptodata/1/paa/US084.COMB.pep:*
9: /cgn2_6/ptodata/1/paa/US085.COMB.pep:*
10: /cgn2_6/ptodata/1/paa/US086.COMB.pep:*
11: /cgn2_6/ptodata/1/paa/US087.COMB.pep:*
12: /cgn2_6/ptodata/1/paa/US088.COMB.pep:*
13: /cgn2_6/ptodata/1/paa/US089.COMB.pep:*
14: /cgn2_6/ptodata/1/paa/US090.COMB.pep:*
15: /cgn2_6/ptodata/1/paa/US091.COMB.pep:*
16: /cgn2_6/ptodata/1/paa/US092.COMB.pep:*
17: /cgn2_6/ptodata/1/paa/US093.COMB.pep:*
18: /cgn2_6/ptodata/1/paa/US094.COMB.pep:*
19: /cgn2_6/ptodata/1/paa/US095.COMB.pep:*
20: /cgn2_6/ptodata/1/paa/US096.COMB.pep:*
21: /cgn2_6/ptodata/1/paa/US097.COMB.pep:*
22: /cgn2_6/ptodata/1/paa/US098.COMB.pep:*
23: /cgn2_6/ptodata/1/paa/US099.COMB.pep:*
24: /cgn2_6/ptodata/1/paa/US100.COMB.pep:*
25: /cgn2_6/ptodata/1/paa/US101.COMB.pep:*
26: /cgn2_6/ptodata/1/paa/US102.COMB.pep:*
27: /cgn2_6/ptodata/1/paa/US60.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2378	100.0	445	22	US-09-890-220-2
2	2285	96.1	440	22	US-09-890-220-5
3	1923	80.9	498	19	US-09-513-996A-69039
4	1796	75.5	367	19	US-09-513-996A-69040
5	1635	68.8	339	19	US-09-513-996A-69041
6	1153.5	48.5	428	26	US-10-219-999-37133

7	1153.5	48.5	428	27	US-60-324-109-23364	Sequence 23364, A
8	923	38.8	611	17	US-09-339-947A-1	Sequence 1, Appl1
9	906.5	38.1	604	17	US-09-339-947A-8	Sequence 8, Appl1
10	833	35.0	481	26	US-10-219-999-48099	Sequence 48099, A
11	833	35.0	481	27	US-60-312-544-10052	Sequence 10052, A
12	654	27.5	813	17	US-09-398-237-2	Sequence 2, Appl1
13	549	23.1	107	22	US-09-890-220-8	Sequence 8, Appl1
14	510.5	21.5	692	25	US-10-177-478-6	Sequence 6, Appl1
15	451.5	19.0	186	22	US-09-880-220-14	Sequence 14, Appl1
16	423	17.8	743	21	US-09-708-427-25830	Sequence 25830, A
17	423	17.8	765	21	US-09-708-427-25829	Sequence 25829, A
18	423	17.8	851	21	US-09-708-427-25828	Sequence 25828, A
19	368.5	15.5	108	22	US-09-890-220-12	Sequence 12, Appl1
20	327	13.8	63	19	US-09-513-996A-69043	Sequence 69043, A
21	324	13.6	145	19	US-09-513-996A-79541	Sequence 79541, A
22	324	13.6	174	19	US-09-513-996A-79540	Sequence 79540, A
23	322.5	13.6	623	21	US-09-708-427-25832	Sequence 25832, A
24	311	13.1	154	22	US-09-890-220-10	Sequence 10, Appl1
25	300	12.6	582	21	US-09-708-427-19473	Sequence 19473, A
26	266	11.2	295	21	US-09-708-427-25824	Sequence 25824, A
27	249	10.5	366	26	US-10-219-999-61010	Sequence 61010, A
28	249	10.5	366	27	US-60-324-109-29363	Sequence 29363, A
29	239	10.1	289	21	US-09-708-427-25825	Sequence 25825, A
30	225	9.5	388	1	PCT-US01-01341-931	Sequence 931, App
31	225	9.5	388	21	US-09-764-864-931	Sequence 931, App
32	225	9.5	388	24	US-10-080-129-931	Sequence 931, App
33	224.5	9.4	302	21	US-09-708-427-19475	Sequence 19475, A
34	224.5	9.4	427	21	US-09-708-427-19474	Sequence 19474, A
35	221	9.3	367	22	US-09-890-220-54	Sequence 54, Appl1
36	221	9.3	739	1	PCT-US01-17936-5	Sequence 5, Appl1
37	221	9.3	739	22	US-09-874-162A-5	Sequence 5, Appl1
38	221	9.3	776	1	PCT-US01-17936-5	Sequence 8, Appl1
39	221	9.3	776	22	US-09-874-162A-8	Sequence 8, Appl1
40	221	9.3	803	22	US-09-890-220-16	Sequence 16, Appl1
41	219	9.2	835	16	US-09-270-767-60845	Sequence 60845, A
42	219	9.2	835	16	US-09-270-849B-191180	Sequence 191180, A
43	219	9.2	947	16	US-09-270-767-45347	Sequence 45347, A
44	219	9.2	947	27	US-60-167-324-1120	Sequence 1120, Ap
45	219	9.2	955	20	US-09-614-150-41697	Sequence 41697, A

ALIGNMENTS

RESULT 1
US-09-890-220-2
; Sequence 2, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
; TITLE OF INVENTION: vernalisation gene VRN2.
; FILE REFERENCE: Mewburn
; CURRENT APPLICATION NUMBER: US/09/890,220
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana

US-09-890-220-2

Query Match 100.0%; Score 2378; DB 22; Length 445;
Best Local Similarity 100.0%; Pred. No. 1.4e-235;
Matches 445; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCRONCRKSSPEEVISTDENLILYCKPVRLNIFHLRSLGNPSFLPCLNKTIGAKRRR 60
|||||

```

Db      1  MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNXYKIGAKRR 60
QY      61  KSRSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFKLOJPHLMSSDIPEFEK 120
Db      61  KSRSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFKLOJPHLMSSDIPEFEK 120
QY      121  FEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKRORGRNNTRLKYCFLPD 180
Db      121  FEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKRORGRNNTRLKYCFLPD 180
QY      181  SPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIPALAHSSLDAGAVILTSE 240
Db      181  SPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIPALAHSSLDAGAVILTSE 240
QY      241  AVVPATKTRKLSAERSEARSHLLQKROFYHSHRVOPMALQVMSDRDSEDEVDVADF 300
Db      241  AVVPATKTRKLSAERSEARSHLLQKROFYHSHRVOPMALQVMSDRDSEDEVDVADF 300
QY      301  EDROMLDDFVYVNDKDEKOPMHLMSFVRKQVYIADGHI SWACEAFSREYERELHRYSLF 360
Db      301  EDROMLDDFVYVNDKDEKOPMHLMSFVRKQVYIADGHI SWACEAFSREYERELHRYSLF 360
QY      361  MCRFLFLIKLMNHGLVDSATINNCTIENCNSSDTTNNNNNSVDRPSDNTNNNTIY 420
Db      361  MCRFLFLIKLMNHGLVDSATINNCTIENCNSSDTTNNNNNSVDRPSDNTNNNTIY 420
QY      421  DHPNDINNNKNNVNDKNNNSRDYIK 445
Db      421  DHPNDINNNKNNVNDKNNNSRDYIK 445

```

RESULT 2

```

US-09-890-220-5
; Sequence 5, Application US/09890220
; GENERAL INFORMATION:
; APPLICANT: Dean, Caroline
; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
; FILE REFERENCE: Newbury
; CURRENT APPLICATION NUMBER: US/09/890,220
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: PCT/GB00/00248
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: GB 9901927.5
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
US-09-890-220-5

```

```

Query Match      96.1%; Score 2285; DB 22; Length 440;
Best Local Similarity 96.8%; Pred. No. 5.6e-226;
Matches 428; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```

```

QY      1  MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNXYKIGAKRR 60
Db      1  MCRONRAKSSPEEVISTDBENLLIYCKPVRLYNIFHLRSLGNPSFLPRCLNXYKIGAKRR 60
QY      61  KSRSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFKLOJPHLMSSDIPEFEK 120
Db      61  KSRSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFKLOJPHLMSSDIPEFEK 120
QY      121  FEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKRORGRNNTRLKYCFLPD 180
Db      121  FEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKRORGRNNTRLKYCFLPD 180
QY      181  SPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIPALAHSSLDAGAVILTSE 240
Db      181  SPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIPALAHSSLDAGAVILTSE 240

```

```

QY      241  AVVPATKTRKLSAERSEARSHLLQKROFYHSHRVOPMALQVMSDRDSEDEVDVADF 300
Db      241  AVVPATKTRKLSAERSEARSHLLQKROFYHSHRVOPMALQVMSDRDSEDEVDVADF 300
QY      301  EDROMLDDFVYVNDKDEKOPMHLMSFVRKQVYIADGHI SWACEAFSREYERELHRYSLF 360
Db      301  EDROMLDDFVYVNDKDEKOPMHLMSFVRKQVYIADGHI SWACEAFSREYERELHRYSLF 360
QY      361  MCRFLFLIKLMNHGLVDSATINNCTIENCNSSDTTNNNNNSVDRPSDNTNNNTIY 420
Db      361  MCRFLFLIKLMNHGLVDSATINNCTIENCNSSDTTNNNNNSVDRPSDNTNNNTIY 420
QY      421  DHPNDINNNKNNVNDKNNNSRDYIK 442
Db      421  DHPNDINNNKNNVNDKNNNSRDYIK 440

```

RESULT 3

```

US-09-513-996A-69039
; Sequence 69039, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513,996A
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69039
; LENGTH: 498
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..498
; OTHER INFORMATION: any n or xaa = unknown
; OTHER INFORMATION: Location 1..498 / Ceres Seq. ID 2185973
US-09-513-996A-69039

```

```

Query Match      80.9%; Score 1923; DB 19; Length 498;
Best Local Similarity 89.4%; Pred. No. 1.6e-188;
Matches 370; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```

```

QY      43  PSFLPRCLNXYKIGAKRRKRSSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFK 102
Db      108  PSFLPRCLNXYKIGAKRRKRSSTGMVFNKYKDCNNNTLOKTEVREDCSCPCSMGCGSFK 167
QY      103  LOFHLSHDLFEFEKFEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKROR 162
Db      168  LOFHLSHDLFEFEKFEETQYVNVSVKLNSTIFEEBGSDDKFEFSLCSKPRKROR 204
QY      163  RGRNNTRLKYCFLPDSPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIP 222
Db      205  RGRNNTRLKYCFLPDSPBLTNGTENGITLLDNGNRGLGYPRATLACQFEMTSNIP 264
QY      223  ALAHSSLDAGAVILTSEAVVPATKTRKLSAERSEARSHLLQKROFYH 271
Db      265  ALAHSSLDAGAVILTSEAVVPATKTRKLSAERSEARSHLLQKROFYH 324
QY      272  SHRVOPMALQVMSDRDSEDEVDVADFEDROMLDDFVYVNDKDEKOPMHLMSFVRKOR 331
Db      325  SHRVOPMALQVMSDRDSEDEVDVADFEDROMLDDFVYVNDKDEKOPMHLMSFVRKOR 384
QY      332  VIADGHI SWACEAFSREYERELHRYSLF MCRFLFLIKLMNHGLVDSATINNCTIENC 391
Db      385  VIADGHI SWACEAFSREYERELHRYSLF MCRFLFLIKLMNHGLVDSATINNCTIENC 444
QY      392  RNSSDTTTNNNNNSVDRPSDNTNNNTIYDHPNDINNNKNNVNDKNNNSRDYIK 445
Db      445  RNSSDTTTNNNNNSVDRPSDNTNNNTIYDHPNDINNNKNNVNDKNNNSRDYIK 498

```

RESULT 4

```

US-09-513-996A-69040
; Sequence 69040, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69040
; LENGTH: 367
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 1..367
; OTHER INFORMATION: any n or xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..367 / Ceres Seq. ID 2185974
US-09-513-996A-69040

```

```

Query Match          75.5%; Score 1796; DB 19; Length 367;
Best Local Similarity 88.7%; Pred. No. 1.3e-175;
Matches 346; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```

```

QY 67 MVENYKDCNNTLOKTEVREDSCPCFSCMLCGSEKGLQPHLNSHDLFEFEKLEEEYQT 126
DB 1 MVEFYKDCNNTLOKTEVREDSCPCFSCMLCGSEKGLQPHLNSHDLFEFEKLEEEYQT 56
QY 127 VNVSKLNSFIEEESGDDDKFEFPLSCSKPRKRRGRNTRRLKTCFLPLDPSLTN 186
DB 57 VNPITMFC-----SKPRKRRGRNTRRLKTCFLPLDPSLTN 97
QY 187 GTENGITLLDNGNRGLGYPEATELAGOFEMTSNIPPAIAHSSLDGAKVILTSEAVPAT 246
DB 98 GTENGITLLDNGNRGLGYPEATELAGOFEMTSNIPPAIAHSSLDGAKVILTSEAVPAT 157
QY 247 KTRKLSAERSSEAR-----SHLLQKQFYHSHRVOPMALEQVMSDRSDSEYVD 295
DB 158 KTRKLSAERSSEARSTAILTERISSSHLLQKQFYHSHRVOPMALEQVMSDRSDSEYVD 217
QY 296 DVADPEDRQMLDDFYDVKDEKQFMHLMNSFVRKORVADGHTISACAFSFEYKEELHR 355
DB 218 DVADPEDRQMLDDFYDVKDEKQFMHLMNSFVRKORVADGHTISACAFSFEYKEELHR 277
QY 356 YSLEFCWRLFLIKLMNHGLVDSATINNNTILFENCRRSSDTTNNNSVDRPDSNTN 415
DB 278 YSLEFCWRLFLIKLMNHGLVDSATINNNTILFENCRRSSDTTNNNSVDRPDSNTN 337
QY 416 NNNIYDHPNDINKNNVNDKNNNSDKYIK 445
DB 338 NNNIYSHPDINKNNVNDKNNNSDKYIK 367

```

RESULT 5

```

US-09-513-996A-69041
; Sequence 69041, Application US/09513996A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
; FILE REFERENCE: 2750-709P
; CURRENT APPLICATION NUMBER: US/09/513, 996A
; NUMBER OF SEQ ID NOS: 81028
; SEQ ID NO 69041
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:

```

```

; NAME/KEY: UNSURE
; LOCATION: 1..339
; OTHER INFORMATION: any n or xaa = unknown
; FEATURE:
; OTHER INFORMATION: Location 1..339 / Ceres Seq. ID 2185975
US-09-513-996A-69041

```

```

Query Match          68.8%; Score 1635; DB 19; Length 339;
Best Local Similarity 87.8%; Pred. No. 4.6e-159;
Matches 318; Conservative 3; Mismatches 7; Indels 34; Gaps 3;

```

```

QY 95 MCGSFKGLQFHLNSHDLFEFEKLEEEYQTVNVSKLNSFIEEESGDDDKFEFSLC 154
DB 1 MCGSFKGLQFHLNSHDLFEFEKLEEEYQTVNVSKLNSFIEEESGDDDKFEFSLC 37
QY 155 SKPRRRRGRGRRNTRRLKVCFLPLDPSLTNLTENGITLLDNGNRGLGYPEATELAGOF 214
DB 38 SKPRRRRGRGRRNTRRLKVCFLPLDPSLTNLTENGITLLDNGNRGLGYPEATELAGOF 97
QY 215 EMTSNIPPAIAHSSLDGAKVILTSEAVPATKTRKLSAERSSEAR-----SHLL 263
DB 98 EMTSNIPPAIAHSSLDGAKVILTSEAVPATKTRKLSAERSSEARSTAILTERISSSHLL 157
QY 264 LQKQFYHSHRVOPMALEQVMSDRSDSEYVDVADFEEDRQMLDDFYDVKDEKQFMHLM 323
DB 158 LQKQFYHSHRVOPMALEQVMSDRSDSEYVDVADFEEDRQMLDDFYDVKDEKQFMHLM 217
QY 334 NSFVRKORVADGHTISACAFSFEYKEELHRYSSLEFCWRLFLIKLMNHGLVDSATINN 383
DB 218 NSFVRKORVADGHTISACAFSFEYKEELHRYSSLEFCWRLFLIKLMNHGLVDSATINN 277
QY 384 CNTILENCRSSDTTNNNSVDRPDSNTNNNIYHPNDINKNNVNDKNNNSDKY 443
DB 278 CNTILENCRSSDTTNNNSVDRPDSNTNNNIYHPNDINKNNVNDKNNNSDKY 337
QY 444 IK 445
DB 338 IK 339

```

RESULT 6

```

US-10-219-999-37133
; Sequence 37133, Application US/10219999
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Edgerton, Michael D
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Liu, Jindong
; APPLICANT: Stein, Joshua
; TITLE OF INVENTION: cDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-10(52726)C
; CURRENT APPLICATION NUMBER: US/10/219, 999
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/324,109
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/312,544
; PRIOR FILING DATE: 2001-08-15
; NUMBER OF SEQ ID NOS: 63520
; SEQ ID NO 37133
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Glycine max
US-10-219-999-37133

```

```

Query Match          48.5%; Score 1153.5; DB 26; Length 428;
Best Local Similarity 54.0%; Pred. No. 3.3e-109;
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;

```

```

QY 1 MCRONCRACKSPPEEYISDENILLYCKPVRLYNIFHLRSLGNPSFLPCLNFKYIGAKRR 60
DB 1 MCRONSPVHHAGEEIEADESLIYCKVELEYINILYRALQNPSPFLRCLRYKIRASRR 60

```

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapm

Page 4

QY	61	KSSSTGVVNNYKDCNNNTLOKTEVREDCSPCCSMLTSCSPFGLOFHLNNSHLLDFEPEFKL	120
Db	61	RLBA-GVIEFNYADRNILIRKTEVEDEFCPCFLMQGCFNGFLAFHLCSHDLFNEFEW	119
QY	121	FEBOYAVNVSKNRSIFEEEGSDD--DKPEPFLCCKPRKROGRNNTRLYKCFL	177
Db	120	TEDYQAVNVSKID--ILRSNVADVYIPQSTFEFCRPRKRRKMSYQIERTANKFL	177
QY	178	PLDSPFLNTENGITLND-----GNGLCYPRATEL	210
Db	178	ELDSP---EGINHGFLOKDDILSCGENVSRSTSEKIFPSSGKNOGCKRPHDPTMDN	234
QY	211	AGCFEMTNSNPP--AIAHSLDAG-ANVYLTSEVNAVYATRTRLSABREANSHELLOK	266
Db	235	LEHVSEFNIIPGVSTAMPDSSVDPCCSISICKSDPALPA-KTKRLSDMRSORRRMLLOK	293
QY	267	KQYHSHRRVQPALAEVMSDRSEDEVDVDADEFDROMLDDVYVANKDEKOFHMLNSF	326
Db	294	RLFHSRRVQPALAEVYLSDRSDEVEDDIDLDEDRMLDDPVDVSKDEQOLHMLNSF	353
QY	327	VKKQRYIADGHSIWACEAFSPRYEKELHRSYSLFPCWRLFLIKILNNGLVDSATINNCT	386
Db	354	MKRORYIADGSHVMACEAFSKLHGKELISPALEFCWRLFMIKILNNGLLDACTMNCSI	413
QY	387	ILENCNNSDPTTNN	401
Db	414	VLDSTYRNEGSGTCKN	428

```

RESULT 7
US-60-324-109-23364
Sequence 23364, Application US/60324109
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovalic, David R.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: DNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10521261B
CURRENT APPLICATION NUMBER: US/60/324, 109
CURRENT FILING DATE: 2001-09-21
NUMBER OF SEQ ID NOS: 33196
SEQ ID NO 23364
LENGTH: 428
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
US-60-324-109-23364

```

Query Match	58.2%	Score 1153.5	DB 27	Length 428
Best Local Similarity	54.0%	Pred. 3.3e-109		
Matches	235	Conservative	57	Mismatches 102
			Indels 41	Gaps 8
QY	1	MCRONRAKSSSEEEYSTDENLLIYCPRVRLRYNLFHRSRGNSFPLPRCLANKIGAKRRR	60	
Db	1	MORONSPVHAAEEIEADESLLIYCPRVRLRYNLFHRSRGNSFPLPRCLANKIRKRRR	60	
QY	61	KSKSGMVVFNKDCNNLTQKTEVREKSCSPFCMSLGGSPKGFPLNLSHDLFEFEFFL	120	
Db	61	RLRA-GIVFNFNRDRYNTLRKTEVTEDFSCFCLMCGNKGKRFHLCSHDLFNFEEVW	119	
QY	121	FEFYQVNVSVLNSFFIIEEGSSDD---DKKEPRLSLCKRRKRRQGGNNTRRLKVCFL	177	
Db	120	TEDYQVNVSVKID---ILRSENVADGVIPQSQFFEFPCSRKRRRRDQVQIKRINVFEL	177	
QY	178	PLDSPSLANGTENGITLLND-----GARGGEYEAREL	210	
Db	178	EIDSP---EGHNGFLQKDDILSCKEGNNVSRTRSSEKIPPSGRNDGGFGFGDHGTMDN	234	
QY	211	AGQFEKTSNPP---AIAHSSLDAG-AKVILTSEAVVPATKRLKLSAERSSEARSHLLQK	266	

D6	235	LEHVESFNIPGVSIAMPOSSVDPECSKICKSDPALPA - RTKKLSMDRSDSRMRMLQK	283
OY	267	ROFYHSHRYQPMALAEVOMSDRSDSEDDVDVADAFEDRMLDMDPDVYKKEDEFOHNLNPF	326
D6	294	RLFHSHRYQPMALAEVULSDRSDSEDDVDIDTDLDRMLDMDPDVDSKDEQOLHNLNPF	353
OY	327	VRFQRYIADSHMACAEAFSRFEKELHNYSLFECWMLFTIKLNNGLVDSATINNCF	386
D6	354	MRRQRYIADSHVPMACAEAFSKLGRLELLSPALFECWMLFTIKLNNGLDADCTINNCSI	413
OY	387	ILENCNSSDTTTN	401
D6	414	VLDSTYNNEGSGTRKN	428

```

RESULT 8
US-09-339-947A-1
? Sequence 1, Application US/09339947A
? GENERAL INFORMATION:
? APPLICANT: YOSHIDA, Nobumasa
? APPLICANT: YAMAI, Yukihito
? APPLICANT: KATO, Yoshihiro
? APPLICANT: HIRATSUKA, Junzo
? APPLICANT: TAKAHASHI, Shigeru
? APPLICANT: MIWA, Tatsushi
? TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
? TITLE OF INVENTION: FLOWERING
? FILE REFERENCE: 032735-001
? CURRENT APPLICATION NUMBER: US/09/339,947A
? CURRENT FILING DATE: 1999-06-25
? PRIOR APPLICATION NUMBER: JP 1998-180065
? PRIOR FILING DATE: 1998-06-26
? PRIOR APPLICATION NUMBER: JP 1999-179043
? PRIOR FILING DATE: 1999-06-24
? NUMBER OF SEQ ID NOS: 20
? SOFTWARE: Patentin Ver. 2.0
? SEQ ID NO 1
? LENGTH: 611
? TYPE: PRT
? ORGANISM: Arabidopsis thaliana
? FEATURE:
? NAME/KEY: ZN_FING
? LOCATION: (306)..(327)
? US-09-339-947A-1

```

[illegible]

QY 145 -DDKEPESLCKP-RKRRGGRNTRRLKVCFLPLDPSLTNG-----TENGITLLN 196
 DB 361 VDPKQOTFEFFSSKKRRRRQKQVSSRQ-----CPHGLGCEVLDKTDHDAVSNS 411
 QY 197 DGNR---GIGYPEATELAGOFEPMTSNIP---PALAHSSLDAGAKVILTSEAVPATKTR 249
 DB 412 EKSRIIPGKHV-----ERIGGAESGGRVPGTSPADVQSCGDDPYQSIAGSTMLOFAKTR 467
 QY 250 KLSARSEARSHLLQKQFYHSHRVOPALAEQVSDRDEDEVDVDDAFEDRQMLDPE 309
 DB 468 KLSIRSDLRNLSLQKQFHFHSHRAQPALEQVLSDRSEDEVDVDDAFEDRQMLDPE 527
 QY 310 VDNDEKQFMHLMNSFVRKQRYIADGHISMACEAFSREYKELHRYSLFMCWRLFLIK 369
 DB 528 VDVTEDEKQMMHNSFVRKQRYIADGHISMACEAFSRLHGPIMVTRPHLMCMRYEMVK 587
 QY 370 LMNGLVDSATINNCNTILE 389
 DB 588 LMNGLIDARFTMNCNTILE 607

RESULT 9
 US-09-339-947A-8
 ; Sequence 8, Application US/09339947A
 ; GENERAL INFORMATION:
 ; APPLICANT: YOSHIDA, Nobumasa
 ; APPLICANT: YANAI, Yukihito
 ; APPLICANT: KATO, Yoshihito
 ; APPLICANT: HIRAKISUKA, Junzo
 ; APPLICANT: TAKAHASHI, Shigeru
 ; APPLICANT: MIWA, Tatsushi
 ; TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
 ; FILE REFERENCE: 032735-001
 ; CURRENT APPLICATION NUMBER: US/09/339, 947A
 ; CURRENT FILING DATE: 1999-06-25
 ; PRIOR APPLICATION NUMBER: JP 1998-180065
 ; PRIOR FILING DATE: 1998-06-26
 ; PRIOR APPLICATION NUMBER: JP 1999-179043
 ; PRIOR FILING DATE: 1999-06-24
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 604
 ; TYPE: PRF
 ; ORGANISM: Oryza sativa
 ; FEATURE:
 ; NAME/KEY: ZN_FING
 ; LOCATION: (310)..(335)
 US-09-339-947A-8

Query Match 38.1%; Score 906.5; DB 17; Length 604;
 Best Local Similarity 33.3%; Pred. No. 1.8e-83;
 Matches 209; Conservative 59; Mismatches 109; Indels 251; Gaps 7;
 QY 1 MCRORCAKSPDEEYISTDENLTIYCKPRLYNIFHLRSLGNSPLPCLNTKIGAKRR 60
 DB 1 MCRORPARLSPDEOLAESEPALYCKPVELYNIIRORSIKNPAFLQRLCLYIARRKK 60
 QY 61 KSRST----- 65
 DB 61 RSLITISLGGTNKELRAQNIPIYVYLARPTNNVSLSEHSPDIYRSRACILTSFEFGN 120
 QY 66 ----- 65
 DB 121 KDYTEATFVIPVKNLATSRACSNLIIISGRABQTFDDNNGSGNHVSGSTLOKLEKRC 180
 QY 66 ----- 65
 DB 181 FNGKIPIDILASSLGNVSLIGHVEMSVTEMTSPFLPEKFLLEDSCSLTFCSSQKVDAT 240
 QY 66 -----GMVFNKDCNNTLQ 80
 I : I I I I I I I I I I

DB 241 GSFOLVISAQAEAGAKDSESPSYSYNDVPSSLTHIIRLSGVLFNKYYNTMQ 300
 QY 81 KTEVEDCSCPCSMICGSRKGLQRLNLSHDLFEFEFLFBEYQVNVSVYLNSTIPE- 139
 DB 301 KTEVEDCSCPCPLVPCGSEFKLGGHILNLSHLEFEFVISECAVNVSLTDSRTSL 360
 QY 140 -EEGSDDDKFEFESLCSKPRKR-RORGNNTRRLKVCFLPLDSP-SLTNGTNGITLLN 196
 DB 361 LAEGV-DPRHQFYSRSRRKKRKYEISDKIRHVPHLYVDSGSPEDAQSGEDDVORE 419
 QY 197 DGNRLGYPEATELAGOFEPMTSNIPALAHSSLDAGAKV---LTSEAVPATKTRKLSA 253
 DB 420 NGS-----SVAHASVDPANSLHGSNLSAPVTLQPGTKRLKSV 456
 QY 254 ERSSEARSHLLQKQFYHSHRVOPALAEQVSDRDEDEVDVDDAFEDRQMLDPEVYN 313
 DB 457 ERADPRNROLQKQFHFHSHRAQPAWMSVFSDRSEDEVDVDDAFEDRQMLDPEVDT 516
 QY 314 KDEQFMHLMNSFVRKQRYIADGHISMACEAFSREYKELHRYSLFMCWRLFLIKLMN 373
 DB 517 KDEKLIMHNMNSFVRKQRYIADGHISMACEAFSRLHGPIMVTRPHLMCMRYEMVK 576
 QY 374 GLVDSATINNCNTILENCSSDPTTTN 401
 DB 577 SLIDARAMNACNTILEGYNLSSDPRKN 604

RESULT 10
 US-10-219-999-48099
 ; Sequence 48099, Application US/102199999
 ; GENERAL INFORMATION:
 ; APPLICANT: Cao, Yongwei
 ; APPLICANT: Edgerton, Michael D
 ; APPLICANT: Hinkle, Gregory J.
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Stein, Joshua
 ; TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
 ; FILE REFERENCE: 38-10(52726)C
 ; CURRENT APPLICATION NUMBER: US/10/219, 999
 ; CURRENT FILING DATE: 2002-08-15
 ; PRIOR APPLICATION NUMBER: US 60/324,109
 ; PRIOR FILING DATE: 2001-09-21
 ; PRIOR APPLICATION NUMBER: US 60/312,544
 ; PRIOR FILING DATE: 2001-08-15
 ; NUMBER OF SEQ ID NOS: 63520
 ; SEQ ID NO 48099
 ; LENGTH: 481
 ; TYPE: PRF
 ; ORGANISM: Zea mays
 US-10-219-999-48099

Query Match 35.0%; Score 833; DB 26; Length 481;
 Best Local Similarity 48.6%; Pred. No. 4.9e-76;
 Matches 180; Conservative 45; Mismatches 83; Indels 62; Gaps 8;
 QY 43 PSFLPCLNKKYKIGARRKRSRSTGMVYVYKCNNTLQTEVREDCSCPCSMICGSEFG 102
 DB 153 PSFLPEILRL-----AGNVLFNKYKYNNTMORTTEVEDSCPCFLVPCGSEFG 201
 QY 103 LOFHNLSSHDLEFEFEKLFEEYQTVNVSVKLSF--IPEEGSDDDKFPFSLCSKPRKR 160
 DB 202 LGCHLNSHDLEHYTELMISEEQAVNVSLKADAMKTEFAEGV-DPRHQTFYSCSFKRR 260
 QY 161 RORGNNTRRLKVCFLPLDPSLTNGT-----ENGITLLDNGRLGLYP 205
 DB 261 R-----RLFTTAEK--FRHVHPHIMESGSHEDAQAGEDDYVORENGLSVAN----- 305
 QY 206 EATELAGOFEPMTSNIPALAHSSLDAGAKVILTSEAVPATKTRKLSARSEARSHLLQ 265
 DB 306 -----TSVDSQPVHGS-----NLSPPVTLQFGKSRKLSARSDPRNNQLQ 347
 QY 266 KQFYHSHRVOPALAEQVSDRSEDEVDVDDAFEDRQMLDPEVDNDEKQFMHLMNS 325

Fri Jun 20 09:03:40 2003

us-09-890-220-2.ram

Page 6

Db 348 KROFESHRAQPMLEOVLSHDSEDEVDLADIEDRMLDDFVDTKDEKLIMHMNS 407
OY 326 FVRKORYIADGHISMACEAFSRFYKEELHRSYSLFPCWMLFLIKIMNGIYDSTATINNCN 385
Db 408 FVRKORYIADGHIPMACCAFNRCHGQOLVONPALLMGWRFPMIKIMNSLIDARMAACN 467
OY 386 TILGCRNSS 395
Db 468 TILGQYODGS 477

RESULT 11

US-60-312-544-10052
Sequence 10052, Application US/60312544
GENERAL INFORMATION:
APPLICANT: Cao, Yongwei
APPLICANT: Edgerton, Michael D
APPLICANT: Hinkle, Gregory J.
APPLICANT: Kovacic, David K.
APPLICANT: Liu, Jingdong
APPLICANT: Stein, Joshua
TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
FILE REFERENCE: 38-10(52726)A
CURRENT APPLICATION NUMBER: US/60/312,544
CURRENT FILING DATE: 2001-08-15
NUMBER OF SEQ ID NOS: 10730
SEQ ID NO 10052
LENGTH: 481
TYPE: PR
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB3206-226-G12_FLI
US-60-312-544-10052

Query Match 35.0%; Score 833; DB 27; Length 481;
Best Local Similarity 48.6%; Pred. No. 4.9e-76;
Matches 180; Conservative 45; Mismatches 83; Indels 62; Gaps 8;

OY 43 PSFLPCINIKIGAKRRKRSRSTGMVVFYKDCNLTOKTEVRDCSCFPGSMICSGFKG 102
Db 153 PSFLPEIRLR-----AGNVLEPKYINIMQTEVEDSCFPLVPCGSKG 201
OY 103 LOFLNMSHDLFEFEKLEEEYQYVAVSYKLSNF--IPEEGSDDDKFEPPSLCSPKRR 160
Db 202 LGCHLMSHDLFEFLMISECOAVANSLKADAKTEFVAEGV-DPRHQTFSSYCSRRKR 260
OY 161 KORGGRNTRRLKVCFLPLDSPSLTNGT-----ENGITLLDNGRGLGYP 205
Db 261 R---RLETTAEK--FRHVPHIMESGSHEDAOAGSEDDVYOREGLSVAN----- 305
OY 206 EATELAGOEFMTSNIPPAIASHSLDGAKVILTSEAVVPATKTRKLSAERSEARSHLLQ 265
Db 306 -----TSVDISOPIVGS-----NLSPPTVQIFGSKRSLSEERDPPNRDLO 347
OY 266 KROFYSHRYOPALBQVNSDRDSEDEVDVADVEDRQMLDDEFVADKDEKOPMHLMS 325
Db 348 KROFESHRAQPMLEOVLSHDSEDEVDLADIEDRMLDDFVDTKDEKLIMHMNS 407
OY 326 FVRKORYIADGHISMACEAFSRFYKEELHRSYSLFPCWMLFLIKIMNGIYDSTATINNCN 385
Db 408 FVRKORYIADGHIPMACCAFNRCHGQOLVONPALLMGWRFPMIKIMNSLIDARMAACN 467
OY 386 TILGCRNSS 395
Db 468 TILGQYODGS 477

RESULT 12

US-09-398-237-2
Sequence 2, Application US/09398237
GENERAL INFORMATION:
APPLICANT: Commonwealth Scientific and Industrial Research Organisation

TITLE OF INVENTION: Novel method of regulating seed development in plants and
FILE REFERENCE: P:\over\Amc\ausseed.pct
CURRENT APPLICATION NUMBER: US/09/398,237
CURRENT FILING DATE: 1999-09-20
EARLIER APPLICATION NUMBER: US 60/101184
EARLIER FILING DATE: 1998-09-21
EARLIER APPLICATION NUMBER: AU PP6063
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: AU PP6062
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: AU PP6061
EARLIER FILING DATE: 1998-09-22
EARLIER APPLICATION NUMBER: AU P01345
EARLIER FILING DATE: 1999-07-01
EARLIER APPLICATION NUMBER: AU P01346
EARLIER FILING DATE: 1999-07-01
NUMBER OF SEQ ID NOS: 213
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 813
TYPE: PR
ORGANISM: Arabidopsis thaliana
US-09-398-237-2

Query Match 27.5%; Score 654; DB 17; Length 813;
Best Local Similarity 24.2%; Pred. No. 3.3e-57;
Matches 184; Conservative 65; Mismatches 147; Indels 364; Gaps 9;

OY 13 EEVISTDNLLIYCKPVATLVNIFRLSGPFLPCINIKIGAKRRKRSRSTGMVVFY 72
Db 56 DDDVYVDENIIRIKIPAVAVYKLETRKNNPYLRSLKYLIOAKKKKNSGGRIRRY 115
OY 73 KDCNNTLO-KTEVEEDSCFPGSMICSGFGLQPHLNSHDLFEFEKLEEEYQ--TYNV 129
Db 116 RYVSKMKTAKAEVENSCFPLCPGCHGLQLHLKSHDAKFFETFAEDHGPEDVY 175
OY 130 SYKLSNFFEE--EGSDDDKFEPPSLCSPKR-KRORGGRNTRRLKVCFLPLDSPSLTN 186
Db 176 SKSDPIKFGVAKDVGVPOLSPFLPCKRNORRODSSNNVKKLVLMELDDLDLPR 235
OY 187 GTENGITLLDNGN-----RG----- 201
Db 200 -----RG----- 201
OY 206 EATELAGOEFMTSNIPPAIASHSLDGAKVILTSEAVVPATKTRKLSAERSEARSHLLQ 265
Db 306 -----TSVDISOPIVGS-----NLSPPTVQIFGSKRSLSEERDPPNRDLO 347
OY 266 KROFYSHRYOPALBQVNSDRDSEDEVDVADVEDRQMLDDEFVADKDEKOPMHLMS 325
Db 348 KROFESHRAQPMLEOVLSHDSEDEVDLADIEDRMLDDFVDTKDEKLIMHMNS 407
OY 326 FVRKORYIADGHISMACEAFSRFYKEELHRSYSLFPCWMLFLIKIMNGIYDSTATINNCN 385
Db 408 FVRKORYIADGHIPMACCAFNRCHGQOLVONPALLMGWRFPMIKIMNSLIDARMAACN 467
OY 386 TILGCRNSS 395
Db 468 TILGQYODGS 477

Qy 319 FMHLNNSFVRKQRYADGHISWACFAFSREYKELHRYSSLFMCWRLFLIKLNMGLVDS 378
 Db 716 YMLYNIFVRKQRYADGHISWACFAFSREYKELHRYSSLFMCWRLFLIKLNMGLVDS 378
 Qy 379 ATINNCNTILENCRRNSDPTTTNNNSVDRPSDSNTNNNN 418
 Db 776 KTFHCKTWTIL-----LSNSDEAGQFTSGSAANANN 805

RESULT 13
 US-09-890-220-8
 ; Sequence 8, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 8
 ; LENGTH: 107
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-8

Query Match 23.1%; Score 549; DB 22; Length 107;
 Best Local Similarity 99.0%; Pred. No. 9.5e-48;
 Matches 100; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MCRONCRASSPEEYISTDENLILYCKPVLNIFHLISLGNPSFLPCLNKYIGAKRRR 60
 Db 1 MCRONCRASSPEEYISTDENLILYCKPVLNIFHLISLGNPSFLPCLNKYIGAKRRR 60

Qy 61 KSRSTGAVYFNKDCNNLTQKTEVREDCSCPCSMGCSFK 101
 Db 61 KSRSTGAVYFNKDCNNLTQKTEVREDCSCPCSMGCSFK 101

RESULT 14
 US-10-177-478-6
 ; Sequence 6, Application US/10177478
 ; GENERAL INFORMATION:
 ; APPLICANT: Dang, Van-Dinh
 ; APPLICANT: Okamoto, Jack
 ; TITLE OF INVENTION: Chimeric Histone Acetyltransferase
 ; FILE REFERENCE: Polypeptides
 ; CURRENT FILING DATE: 2002-06-21
 ; PRIOR FILING DATE: 2001-06-22
 ; PRIOR APPLICATION NUMBER: U.S. 60/300,135
 ; NUMBER OF SEQ ID NOS: 45
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 692
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-10-177-478-6

Query Match 21.5%; Score 510.5; DB 25; Length 692;
 Best Local Similarity 22.3%; Pred. No. 1.7e-42;
 Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;

Qy 81 KTEVREDCSCPCSMGCSFKGLQFHLNNSHDLFEFFKLEEXYO-TVNVSVKLNSEIF 138
 Db 4 KAEVVENSCPCFLIPCGHGGLQHLNNSHDLFEFFKLEEXYO-TVNVSVKLNSEIF 138

Qy 139 E--EEGDDDKFEPELSCKPR-KRROGRNRTRLKVCFLPLDSPSLTNGTENGITLL 195
 Db 64 GYLKDVGNPOLSPFLFOSKRNRRQDDSNVKKLVNLMELDLDDLPSTENDSTHV 123
 Qy 196 NDN-----RG----- 199
 Db 124 NDDNVSPBRAHSEKISDILTTQLAIAESSEPKVPHVNDGNVSPBRAHSAKNEST 183
 Qy 200 ----- 199
 Db 184 HVNDDDVSPBRAHSEKNESTHVNEEDNISPPRAHSSKNESTHNEDEVSPPRTS 243
 Qy 200 -----RG----- 201
 Db 244 SKETSDILTTQPAIVESEPKVRKGRKRLYAKRYARETQPAIAESSEPKYLHVDE 303
 Qy 202 --LGYPEATELAGOEMTSNIPPAIAHSS----- 228
 Db 304 NVSPPEAHSLKASDILTTQPAIAESSEPKVPHVNDGNVSPBRAHSSKKNSTKRV 363
 Qy 229 ----- 228
 Db 364 DNVSPPKTRSSKTSIDILTTQPTIAESSEPKVPHVNDGNVSPBRAHSSKKNSTKRN 423
 Qy 229 -----LDAGAVIILSEAVP----- 244
 Db 424 DNDIPSPKTRSSKTSNLTFTQPAIAESSEPKVPHVNDGNVSPBRAHSSKKNSTHKK 483
 Qy 245 -----ATKRLSAERSEARSHLLQKR 267
 Db 484 DNASLPKTRSSKTSIDILTTQPAIAESSEPKVPHVNDGNVSPBRAHSSKKNSTHKK 543
 Qy 268 QETHSHRYQPALEOVMSDRDSEDDVADFEQMLDDEFVDVNDKDEQPMHLMNSFV 327
 Db 544 QFYHSQTPQPMTEFQVMSDESENETDYDALISRLRLERLYGVSKERKRYMYLNTFV 603
 Qy 328 KRQRYIADGHISWACFAFSREYKELHRYSSLFMCWRLFLIKLNMGLVDSATINNCNTI 387
 Db 604 KRQRYIADGHISWACFAFSREYKELHRYSSLFMCWRLFLIKLNMGLVDSATINNCNTI 387
 Qy 388 LENCRRNSDPTTTNNNSVDRPSDSNTNNNN 418
 Db 664 L-----LSNSDEAGQFTSGSAANANN 684

RESULT 15
 US-09-890-220-14
 ; Sequence 14, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Mewburn
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 1999-01-28
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 186
 ; TYPE: PRT
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-14

Query Match 19.0%; Score 451.5; DB 22; Length 186;
 Best Local Similarity 50.8%; Pred. No. 2.6e-37;
 Matches 99; Conservative 32; Mismatches 43; Indels 21; Gaps 7;

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapm

Page 8

```
QY 237 LTRAVVPATKTRKLSERSEARSHLL-LQKROFYHSRVPMALEQVMSDRDSEDEVD- 294
DB 6 LITRAKVP-----AKRSKATSHYLPJHKROFYHSRTGQPLSLQVMSDRDSEDEVDK 57
QY 295 -DVADEPDRQMLDDFVDVVK-DEKQFMHLMNSFVKQRYADGHI SWACEAFSRYEKE 352
DB 58 NDDAHLLESQMLNGSDEDEVEI AERFIKLMNSFWKQRYADGHI PWACEAFSRLHJQE 117
QY 353 LHRYSLEWCMWRLFIKLMNHGLVDSATJNNCNTILLENCNRSDDTTTNNNSYDRPDS 412
DB 118 LRSNLSLJLCWROFMIRQMDYGLDRVTMKNCKNTIYH--NISTNDINNNT-----R 170
QY 413 NTNNNNIVDHPNDIN 427
DB 171 TTDNMDVVD--DDIN 183
```

Search completed: June 13, 2003, 15:43:51
Job time : 313 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein -- protein search, using sw model

Run on: June 13, 2003, 15:30:54 ; Search time 54 Seconds
(without alignments) 2102.183 Million cell updates/sec

File: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRQNCRAKSSPEEVISTDE.....INKNKNVNDKDNRSRDYIK 445

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1155225 seqs, 255096386 residues

Total number of hits satisfying chosen parameters: 1155225

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/1/paa/PCF_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US10_NEW_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1153.5	48.5	428	6	US-10-424-599-254264
2	1153.5	48.5	435	6	US-10-425-114-46823
3	833	35.0	481	6	US-10-425-114-54921
4	654	27.5	813	6	US-10-231-778-2
5	648.5	27.3	640	6	US-10-437-963-193525
6	617.5	26.0	201	6	US-10-437-963-114092
7	575.5	24.2	242	6	US-10-424-599-283985
8	510.5	21.5	692	6	US-10-198-723-6
9	491	20.6	166	6	US-10-424-599-254263
10	269	11.3	102	6	US-10-424-599-183487
11	256.5	10.8	116	6	US-10-424-599-216286
12	255.5	10.7	151	6	US-10-437-963-199908
13	249	10.5	394	6	US-10-425-114-70148
14	190	8.0	263	6	US-10-437-963-184120
15	159	6.7	195	6	US-10-424-599-250979
16	140.5	5.9	116	6	US-10-424-599-152510
17	132	5.6	749	6	US-10-369-493-21933
18	128	5.4	106	6	US-10-437-963-116134
19	124	5.2	1288	1	PCF-US02-38360-2
20	122	5.1	86	6	US-10-437-963-201278
21	122	5.1	123	6	US-10-437-963-109658
22	119	5.0	1094	1	PCF-US01-29288A-216
23	119	5.0	1094	6	US-10-149-310-216
24	118	5.0	249	6	US-10-437-963-135834
25	115.5	4.9	2184	6	US-10-304-093-6
26	114.5	4.8	879	6	US-10-282-122A-52571

27	114	4.8	964	1	PCF-US01-29288A-256	Sequence 256, App
28	114	4.8	964	6	US-10-149-310-256	Sequence 256, App
29	113.5	4.8	105	6	US-10-437-963-112367	Sequence 112367, App
30	113.5	4.8	654	5	US-09-134-000C-6645	Sequence 6645, App
31	113.5	4.8	654	5	US-09-134-000C-6645	Sequence 6645, App
32	113.5	4.8	654	5	US-10-434-665-6645	Sequence 6645, App
33	112	4.7	104	6	US-10-437-963-198200	Sequence 198200, App
34	108.5	4.6	905	6	US-10-369-493-21936	Sequence 21936, App
35	108	4.5	336	6	US-10-424-599-223789	Sequence 223789, App
36	108	4.5	336	6	US-10-438-246-32035	Sequence 32035, App
37	108	4.5	337	6	US-10-425-114-39945	Sequence 39945, App
38	107.5	4.5	192	6	US-10-437-963-152724	Sequence 152724, App
39	107.5	4.5	833	6	US-10-369-493-21936	Sequence 21936, App
40	107	4.5	138	6	US-10-437-963-176416	Sequence 176416, App
41	107	4.5	313	6	US-10-425-114-52231	Sequence 52231, App
42	106.5	4.5	833	6	US-10-310-154-480	Sequence 480, App
43	106	4.5	328	6	US-10-437-963-107965	Sequence 107965, App
44	105.5	4.4	787	6	US-10-282-122A-72461	Sequence 72461, App
45	104.5	4.4	97	6	US-10-271-869-77	Sequence 77, App

ALIGNMENTS

RESULT 1	US-10-424-599-254264	Sequence 254264, Application US/10424599
GENERAL INFORMATION:		
APPLICANT:	La Rosa Thomas J	
APPLICANT:	Kovalic David K	
APPLICANT:	Zhou Yihua	
APPLICANT:	Cao Yongwei	
TITLE OF INVENTION:	Soy Nucleic Acid Molecules and Other Molecules Associated with	
TITLE OF INVENTION:	Plants and Uses Thereof for Plant Improvement	
FILE REFERENCE:	38-21(53223)B	
CURRENT APPLICATION NUMBER:	US/10/424,599	
CURRENT FILING DATE:	2003-04-28	
NUMBER OF SEQ ID NOS:	285684	
SEQ ID NO 254264		
LENGTH:	428	
TYPE:	PRT	
ORGANISM:	Glycine max	
FEATURE:		
OTHER INFORMATION:	Clone ID: PAT_MRT3847_71624C.1.pep	
US-10-424-599-254264		
Query Match	48.5%; Score 1153.5; DB 6; Length 428;	
Best Local Similarity	54.0%; Pred. No. 3.9e-100;	
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;		
QY	1 MCRQNCRAKSSPEEVISTDENLLTYCKRYRLNIFHLASLGNPSFLPCINRYKIKARRR 60	
DB	1 MCRQNSPVHAAEEIEADESLTYCKPELYNLIRLQNPFLRCLRYKIRASRR 60	
QY	61 KSRSGMVFVKKDCNNTLQKTEVEDSCPCSMGCSFKGLQFLNLSHDLFFEEKL 120	
DB	61 RLRAGIYFENRDNRYNLIRKTEVEDSCPCMLQCGFKLRHLCSHDLFFEEFV 119	
QY	121 FEETGVNYSVKLNFIFEEBGSDD--DKPEPFLCSKPRRRRORGRNNTRLKVCFL 177	
DB	120 TEDYQAVNVYSKID--ILRSENVADGVIPQSTFFCFSPRRRRKDSVQIKRNIVFL 177	
QY	178 PIDSLSLNGTENGTLND-----GNRGIGYPEATEL 210	
DB	178 EIDSP---EGHNGFLQKDDLLSKGEVNSTSEKIFPSGRNDGKFGFDHGTMDN 234	
QY	211 AGQFEMTSNIP--AIHSSLDAG-AVYLTSEAVVPATKTRKLSARSEARSHLLQK 266	
DB	235 LEHVSSFNIPGVSIAMPOSSVDPSCSKSICKSDPALPA-KTKKLSMDRSDRNRLQK 293	
QY	267 ROFYSHRYQVPALQVMSDRSDEVDVADFFEDRQMLDFVYVKNDEKQFHLNNSF 326	
DB	294 RLFFSHRYQVPALQVMSDRSDEVDVADFFEDRQMLDFVYVKNDEKQFHLNNSF 353	

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapn

Page 2

QY 327 VRKORVADGHISMACEAFSEFEKELHRYSSLFQCMWRLFKIKMNGLDVDSATINNCT 386
DB 354 MKRORVADGHVPMACEAFSEKHLKELISSPALFQCMWRLFKIKMNGLDVDSATINNCT 413
QY 387 ILKCRSSSDTTTN 401
DB 414 VLDSEYRNEGSGTRKN 428

RESULT 2
US-10-425-114-46823
Sequence 46823, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 46823
LENGTH: 435
TYPE: PRN
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 701055296_FLI.pep
US-10-425-114-46823

Query Match 48.5%; Score 1153.5; DB 6; Length 435;
Best Local Similarity 54.0%; Pred. No. 46-100;
Matches 235; Conservative 57; Mismatches 102; Indels 41; Gaps 8;
QY 1 MCRCNCRANSSPEVISTDNLITCKPVRLYNIHLSLGNPSEFLPRCLNYKIGANKRR 60
DB 8 MCRCNCRANSSPEVISTDNLITCKPVRLYNIHLSLGNPSEFLPRCLNYKIGANKRR 67
QY 61 KSRSTGVVFNKDCNNLTOKTEVEDSCSPCSMLGSGFKGLQFHLNLSHDLFEFEFKL 120
DB 68 RLRA-GIVIFNDRNIRLKTVEDEFSQPCMLQCNFKGLFHLNLSHDLFEFEFKL 126
QY 121 FEYQTVNVSVKLNSFIFEEGSD--DKFEPFLSKPKRRKQRRGRRNTRRLKVCFL 177
DB 127 TEDYQAVNVSVKID--ILRSENVADGVIFQSGTFEFCSPRRKRRKDSVQIEKRTNFKL 184
QY 178 PLDSPSLTNGTENGITLND-----GNRGLGYPEATEL 210
DB 185 ELDSP--EGIHNGFLQKDDLLSCGENVSRTSRSEKIFPSGRNDGKFGPDHPTMDN 241
QY 211 AGOFEMTSNIP--AIHSSLDAG-AKYLITSEAVVPATYTRKLSAERSFARSHLLQK 266
DB 242 LEHVESSEFNIPGSIAMPOSSVDPCSKSICKSPALPA-KTKLMSDRSRRRMLQK 300
QY 267 ROFYHSHRVPALAEQVMSDRDSEVDVADPEDROMLDFVDYNDKDEQFHLNLSNF 326
DB 301 RLFEHSHRVPALAEQVMSDRDSEVDVADPEDROMLDFVDYNDKDEQFHLNLSNF 360
QY 327 VRKORVADGHISMACEAFSEFEKELHRYSSLFQCMWRLFKIKMNGLDVDSATINNCT 386
DB 361 MKRORVADGHVPMACEAFSEKHLKELISSPALFQCMWRLFKIKMNGLDVDSATINNCT 420
QY 387 ILKCRSSSDTTTN 401
DB 421 VLDSEYRNEGSGTRKN 435

RESULT 3
US-10-425-114-54921

Sequence 54921, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jindong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 54921
LENGTH: 481
TYPE: PRN
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: UC-ZMFLB73226G12_FLI.pep
US-10-425-114-54921

Query Match 35.0%; Score 833; DB 6; Length 481;
Best Local Similarity 48.6%; Pred. No. 86-70;
Matches 180; Conservative 45; Mismatches 83; Indels 62; Gaps 8;
QY 43 BSFLPRCLNYKIGAKRRKRSSTGVVFNKDCNNLTOKTEVEDSCSPCSMLGSGFKG 102
DB 153 BSFLPRCLNYKIGAKRRKRSSTGVVFNKDCNNLTOKTEVEDSCSPCSMLGSGFKG 201
QY 103 LQFHLNLSHDLFEFEFKLFEYQTVNVSVKLNSF--IFEESGDDDKFEPFLSKPKRR 160
DB 202 LQCHLNLSHDLFEHLMTSECOQAVNVSLLADAMKTEFAEGV-DRHQTFSYCSGRFKR 260
QY 161 RQGRNRRRLKVCFLPLDSPSLTNGT-----ENGITLNDGNGLGYP 205
DB 261 R---RLTETAEK--FRVHPHIMESGSHDAQSGEDDYVQRENGLSVAN----- 305
QY 206 EATELAGOFEMTSNIPALAHSSLDAGAKYLITSEAVVPATYTRKLSAERSFARSHLLQ 265
DB 306 -----TSVDPSPQVHGS-----NLSPPTVLQFGKSRKLSAERSDPRRRLQ 347
QY 266 KROFYHSHRVPALAEQVMSDRDSEVDVADPEDROMLDFVDYNDKDEQFHLNLSNF 325
DB 348 KROFYHSHRVPALAEQVMSDRDSEVDVADPEDROMLDFVDYNDKDEQFHLNLSNF 407
QY 326 FVRKORVADGHISMACEAFSEFEKELHRYSSLFQCMWRLFKIKMNGLDVDSATINNCT 385
DB 408 FVRKORVADGHVPMACEAFSEKHLKELISSPALFQCMWRLFKIKMNGLDVDSATINNCT 467
QY 386 ILKCRSSSDTTTN 401
DB 468 ILKCRSSSDTTTN 435

RESULT 4
US-10-231-778-2
Sequence 2, Application US/10231778
GENERAL INFORMATION:
APPLICANT: Bildeau, Pierre
APPLICANT: Chaudhury, Abdul M.
APPLICANT: Dennis, Elizabeth S.
APPLICANT: Koltunow, Anna M.G.
APPLICANT: Luo, Ming
APPLICANT: Peacock, William J.
TITLE OF INVENTION: Method for inducing seed development by down-regulating
FILE REFERENCE: 72-98A
CURRENT APPLICATION NUMBER: US/10/231,778
CURRENT FILING DATE: 2002-11-08
PRIOR APPLICATION NUMBER: 09/398,237
PRIOR FILING DATE: 1999-09-20
PRIOR APPLICATION NUMBER: 60/101,184

```

QY      13 EEVSTEBENLLIYCKPALVYIIFHLRSIGNSPFRCNTNYIGAKRRKRSSTGAVWNY 72
Db      56 DDVDVDENITIKYIKPAVYIKKLETRSNNNYFHLRSIKYIIOAKKKKSSGCGIRRENY 115
QY      73 KDCNNLTQ-KTEVRDCSCPECSMIGSGFKLOPHLNSHDLFEFEFKLFEYO--TVNY 129
Db      116 RDVSNKMTLKAEEVENFSCPECLIPCGHEGLDHLTKSHDAFKEFEYRAEKDHGPEVDY 175
QY      130 SVKNSLIFE--EESGSDDKFEPSLCSKR-KRRQRCGRNTRLKCYFLDPSFLTN 186
Db      176 SVSDDTKFGYKLDVGNPOLSPILFCSKNRQRRQDSDNNVKLNLVLMELDLDLPR 235
QY      187 GTENGILLNDGN----- 199
Db      236 GTEENDSHVANDNVNSPPRAHSEKISDILTTQOLAIESSPEKYPHVNDGVSPRAH 295
QY      200 ----- 199
Db      296 SSAEKNESTHVNDDDVSSPPRAHSEKNESTHVNEINISSPPKAHSSKNESTHNDSD 355
QY      200 -----RG----- 201
Db      356 VSPFRTRSSKETDILTTQPALYBSEPKVRKGRSKKOLYAKRYKARETOPAIAESE 415
QY      202 -----LGYPEATELAGOPEMTSNIPPAIAHSS----- 228
Db      416 PKVLHVNDENVSSPEAHSLSEKASDILTTQPAIAESEPKYPHVNDENVSTPRAHSSK 475
QY      229 ----- 228
Db      476 KNKSTRKNVDNVPSPKTRSSKKTSDILTTQOPTAESSEPKVRHVNDNVNSTPRAHSS 535
QY      229 -----LDAGAKVILTSEAVP----- 244
Db      536 KKNKSTRKNDNISPSPKTRSSKKTSNLTTQPAIAESEPKYPHVNDKXSTPRAHSS 595
QY      245 -----ATKTRKLSAERSEA 258
Db      596 KKNKSTRKKDDNAGLPPKTRSSKKTSDILATTOPAKAEPSEPKYTVRSRRKELHAERCEA 655
QY      259 RSHLLQKROEYSHRQVPMALDEOVMSRSDSEDEVDVADFEQMDLDDPYDVAKDEQO 318
Db      656 KRLERLGROQYHSQTOPMTEFEOVMSNEDSENEDDVALDISERLRLERLGVASKEER 715
QY      319 FMHLNMFVRRKORVITADGHISWACAEASREYKEKELHRYLSFLFCWCRLLIKLMNGLVDS 378
Db      716 YMYLWNLFVRRQORVITADGHYPMACDEEFAKLKEKKNSSPFDWMYRMRIKLMNNGLICLA 775
QY      379 ATINNCNTLLENCRNSSDTTTTNNNSVYDRPDSNTNNNN 418

```

```

0y      217  ---PALQVNSDRDSEDEVDVDFEDRQMLDFVDVNDKQKQFMHLMNSFVRKQRY 333
        || :||:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
db      513  MSVPATFEFVLSDNSEDEVDVDIADLEDRRLDQFVDVTKDEKRIMHMNSFIKQSL 572

```

OY 334 ADGHSMACEAFSEFYKEKELHRYSSLEFCWMLFLIKLNHGLVDSATINNCNTILENCRN 393
DB 573 ADHVPWACEAFSEFRRHGEELLENLALMGWEMKIKLNHSHLSARIMDTCTNRIIDDKN 632

RESULT 6

US-10-437-963-114092
Sequence 114092, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 114092
LENGTH: 201
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1.pep
US-10-437-963-114092

Query Match 26.0%; Score 617.5; DB 6; Length 201;
Best Local Similarity 62.6%; Pred. No. 5.8e-50;
Matches 114; Conservative 26; Mismatches 39; Indels 3; Gaps 1;

OY 223 AINSSIDAGAKV--ILNSEAVVPATKTRKLSAERSSEARSHLLLOKROFYSHRVPMA 279
DB 20 SVNAAVSVPANSSIHGNSINSAPTVLOFGKTRKLSYERADPRNRKOLLOKROFYSHRVPMA 79
OY 280 LEOVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFQMLNNSFVRKORYADGHIS 339
DB 80 LEOVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFQMLNNSFVRKORYADGHIS 139
OY 340 WACEAFSEFYKEKELHRYSSLEFCWMLFLIKLNHGLVDSATINNCNTILENCRNSSDTT 399
DB 140 WACEAFSEFYKEKELHRYSSLEFCWMLFLIKLNHGLVDSATINNCNTILENCRNSSDPK 159
OY 400 TN 401
DB 200 KN 201

RESULT 7

US-10-424-599-283985
Sequence 283985, Application US/10424599
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 283985
LENGTH: 242
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1.pep
US-10-424-599-283985

Query Match 24.2%; Score 575.5; DB 6; Length 242;
Best Local Similarity 48.2%; Pred. No. 7e-46;
Matches 123; Conservative 38; Mismatches 63; Indels 31; Gaps 6;

OY 145 DKEPEFSLCSKPKRRRORGGRRNTTRKLCFPLDPSLTNGTITLNDGRRGLGY 204
DB 13 DRYVOTFEFCCKPLKRRRTADQS---LKNA-VGLESSPPAGCTD--ILEKDDG----- 59
OY 205 PEATRLAGOFMTSNIPPAIHSSLDGAKVILT--SEAVVPATKTRKLSAERSSEARSHL 262
DB 60 -----ISATIRSRPRDSDVSQMSDDQAVLOFPAKTRKLSIERPPRNS 104
OY 263 LLOKROFYSHRVPMALEQVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFQML 322
DB 105 FLKROFYSHRVPMALEQVMSDRDSEDEVDVADFEEDROMLDFVDVNDKROFQML 164
OY 323 WNSFVRKORYADGHISWACEAFSEFYKEKELHRYSSLEFCWMLFLIKLNHGLVDSATIN 382
DB 165 WNSFVRKORYADGHISWACEAFSEFYKEKELHRYSSLEFCWMLFLIKLNHGLVDSATIN 224
OY 363 NCNTILENC-RNSSD 396
DB 225 DCNTILEYORNSD 239

RESULT 8

US-10-198-723-6
Sequence 6, Application US/10198723
GENERAL INFORMATION:
APPLICANT: Dang, Van-Dinh
APPLICANT: Okamoto, Jack
TITLE OF INVENTION: CHIMERIC POLYPEPTIDES MODULATING
FILE REFERENCE: 11696-043001
CURRENT APPLICATION NUMBER: US/10/198,723
CURRENT FILING DATE: 2002-07-05
PRIOR APPLICATION NUMBER: US 60/303,654
PRIOR FILING DATE: 2001-07-06
NUMBER OF SEQ ID NOS: 100
SOFTWARE: Fastseq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 692
TYPE: PRT
ORGANISM: Arabidopsis thaliana
US-10-198-723-6

Query Match 21.5%; Score 510.5; DB 6; Length 692;
Best Local Similarity 27.3%; Pred. No. 4.1e-39;
Matches 154; Conservative 53; Mismatches 121; Indels 363; Gaps 8;

OY 81 KTEVREDCSCFCMLGSGFKGIQFHNSHDLEFEFFLFEDYQ--TYNVSYKLNSTIF 138
DB 4 KAEVVENSCFCFLIPGHEGLQHLKSSHDNFKFEYRAERDNGPEVDVSVKSDITKF 63
OY 139 E-BEGSDDDKFEPEFSLCSKPKRRRORGGRRNTTRKLCFPLDPSLTNGTITL 195
DB 64 GVLKDDVGNPQLSPFTFCSSKRNROORDSDNNVKLLNLLMELDDLDPRTENDSTHV 123
OY 196 NDGN----- 199
DB 124 NDDVNSPPRAHSEKISDILTTOLAIESSEPKVPVHNDGNVSSPPRAHSSAEKNEST 183
OY 200 ----- 199
DB 184 HVNDDDDVSSPPRAHSEKNESTHVNDNISPPRAHSSKNESTHNDDEVSPPTRS 243
OY 200 -----RG----- 201
DB 244 SKETSDILTTTPAIVEPEPKVRRGSRKOLYAKRYKAROTPAIAESSEPKVLYHND 303
OY 202 --LQVPEATELAGFEWTSNIPPAIHSS----- 228

Db 304 NVSPPEAHSLEKASDILTTTQPAIESSEPKVPHVNDENVSTPRAHSSKKNKSTKRV 363
QY 229 ----- 228
Db 364 DNVSPPKTRSSKTSIDILTTTQPTIAESSEPKVPHVNDENVSTPRAHSSKKNKSTKRN 423
QY 229 -----LDAGARVILITSEAVP----- 244
Db 424 DNIPSPKTRSSKTSNLTFRTPAIAESEPKVPHVNDKVSSTPRAHSSKKNKSTHK 483
QY 245 -----ATKTKLSAENSEASHLLQKR 267
Db 484 DNASLPKTRSSKTSIDILATTPAKAEPSEPKVTRVSRREKLAERCAERLERLKGR 543
QY 268 QFYHSHRYQPMALQOVMSDRDSEDDVDVADFEEDROMLDFVDYKDEKOPMHLMSFV 327
Db 544 QFYHSHRYQPMALQOVMSDRDSEDDVDVADFEEDROMLDFVDYKDEKOPMHLMSFV 603
QY 328 RKQVIADGHSIMWACEAFSRFEKELHRYSLFWCRLFLIKLMHGLVDSATINNCMTI 387
Db 604 RKQVIADGHSIMWACEAFSRFEKELHRYSLFWCRLFLIKLMHGLVDSATINNCMTI 663
QY 388 LENCNNSDITTTNNNNVDRPSDNTNNNN 418
Db 664 L-----LSNDEAGQFTSGSAANANN 684

RESULT 9

US-10-424-599-254263
; Sequence 254263, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 254263
; LENGTH: 166
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(166)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1.pep
US-10-424-599-254263

Query Match 20.6%; Score 491; DB 6; Length 166;
Best Local Similarity 60.2%; Pred. No. 4e-38;
Matches 100; Conservative 21; Mismatches 39; Indels 6; Gaps 3;

QY 1 MCRONCRAKSSPEEYISIDENTILYCKPRLYNIFHLSIGSPSLRPLCNKIGAKRR 60
Db 1 MCRONSPVHHGEEETIADDESLLTYCKPELNYILYRRLOQPSFLRRLCKYKIRARRR 60
QY 61 KSRGTGVVFNKDCNNTLOKTEVEDSCSPFCMISGFKLOFHLNSSHDLFEFEFKL 120
Db 61 RLRA-GIYFNVRDHYNLIRKTEVEDSCSPFCMISGFKLRHLNSSHDLFEFEFKV 119
QY 121 FEEYTVAVSVKLSFIIEEGSD--DKFEPSLCKPKRROR 163
Db 120 TEDYQAVNVSVKIN--IFEENVADGVIPQSOTFFCSPRRNRKR 163

RESULT 10
US-10-424-599-183487
; Sequence 183487, Application US/10424599

; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 183487
; LENGTH: 102
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(102)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_136701C.1.pep
US-10-424-599-183487

Query Match 11.3%; Score 269; DB 6; Length 102;
Best Local Similarity 52.1%; Pred. No. 1.9e-17;
Matches 50; Conservative 14; Mismatches 32; Indels 0; Gaps 0;

QY 289 SEDVDVDDVADFEEDROMLDFVDYKDEKOPMHLMSFVRKQRYIADGHSIMWACEAFSR 348
Db 7 SOQEVDDHDDVEDPATYHLHLAHSDEXKXFMHXCHSLKRRKSVCPHGVWMACEAFKOF 66
QY 349 YEKELHRYSSLFWCRLFLIKLMHGLVDSATINNC 384
Db 67 YEKELLSLALFWRRRLFWIKIKMNYGLHARIMNKC 102

RESULT 11

US-10-424-599-216298
; Sequence 216298, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216298
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37346C.1.pep
US-10-424-599-216298

Query Match 10.8%; Score 256.5; DB 6; Length 116;
Best Local Similarity 41.2%; Pred. No. 3.4e-16;
Matches 47; Conservative 26; Mismatches 38; Indels 3; Gaps 2;

QY 293 VDDVADFEEDROMLDFVDYKDEKOPMHLMSFVRKO--RVYADGHSIMWACEAFSRFE 350
Db 1 MNEGAQELTDRRL-AILDAPHYAEFLSLMAAFIKKHGRVLADGHIMWACEAFRTKYS 59
QY 351 KELHRYSSLFWCRLFLIKLMHGLVDSATINNCYTIENCNNSDITTTNNNN 404
Db 60 AFEAQSNSLAWMMRWEMIKLINQGLTKSTFYACNNILIEQCKRQNDPQRMQD 113

RESULT 12
US-10-437-963-199908

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rapn

Page 6

```

? Sequence 199908, Application US/10437963
? GENERAL INFORMATION:
? APPLICANT: La Rosa, Thomas J.
? APPLICANT: Kovalic, David K.
? APPLICANT: Zhou, Yihua
? APPLICANT: Cao, Yongwei
? APPLICANT: Wu, Wei
? APPLICANT: Boukharov, Andrey A.
? APPLICANT: Barbazuk, Brad
? APPLICANT: Li, Ping
? TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
? TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
? FILE REFERENCE: 38-21(53221)B
? CURRENT APPLICATION NUMBER: US/10/437,963
? CURRENT FILING DATE: 2003-05-14
? NUMBER OF SEQ ID NOS: 204966
? SEQ ID NO 199908
? LENGTH: 151
? TYPE: prt
? ORGANISM: Oryza sativa
? FEATURE:
? OTHER INFORMATION: Clone ID: PAT_MRT4530.9542C.1.pep
? US-10-437-963-199908

```

	Query Match	10.7%	Score 255.5;	DB 6	Length 151;
	Best Local Similarity	42.4%;	Pred. No. 6.1e-16;		
	Matches 67;	Conservative 15;	Mismatches 39;	Indels 37;	Gaps 5
Oy	4 ONCRAK----	SPEEYISIDENLLICRQVRVRLYNFHEHRSIGNSFLPRCLNXYMGKKRR	60		
	:	:			
Db	25 QGACADMSSESYSYSYND-----	VPSLSLTHIRRS-----	58		
Oy	61 KKSSTMMVPNKDKDNNTLOKTEVREDCSGFCMLCGSRGTGFPHNSHDLEFEFFLT	120			
	:	:			
Db	59 -----GNVFENKIYNNHTOKTEVTEDESCFCLPCCGSRGJACHLNASHDLFIHFMTI	113			
Oy	121 FEERYQVANNYSVALNGSFTIE--EESGDDDKFEPSLCSK	156			
	:	:			
Db	114 SEECQAIVNYSLKTDISMRTLLAEGV-DPRDHTFSYNR	150			

```

RESULT 13
US-10-425-114-70148
: Sequence 70148: Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jindong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kowaljc, David K.
: APPLICANT: Steven E.
: APPLICANT: Nabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
: TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
: FILE REFERENCE: 38-21(531)B
: CURRENT APPLICATION NUMBER: US/10/425,114
: NUMBER FILING DATE: 2003-04-28
: NUMBER OF SEQ. ID NOS: 73128
: SEQ ID NO 70148
: LENGTH: 394
: TYPE: PRT
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-2MFLB73067D07_FLI pep
: US-10-425-114-70148

```

```

Query Match      10.5%  Score 249; DB 6;  Length 394;
Best Local Similarity  0.8%  Pred. No. 9, Se-15;
Matches          70;  Conservative 17;  Mismatches 29;  Indels 220;  Gaps 2;

OY      3  RQNCRAKSSPEEVIDENLLTYCKPVALRNIHFILNSLGNPSFLPCLCAVYTKGAKRR-- 60
          |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      55  RQQLRALSPEDKRLAENDLALCYCKPVELNLIQRRAMKNPLIQICLCLLYNHARKKRI 114

```

QY	61	-----	-----	60
Db	115	QTITISGSGTNTELOTHYVPEPLYLLARPTSNLSLEGHSPIYRFRSHVCLTFSSEGNKD	174	
QY	61	-----KSRST-----	65	
Db	175	NSEATFTIPDYKSLSTSRACNHDIIIFISCGQGSQNGEDNCSGNHVEDSLQMLEGKCSW	234	
QY	66	-----	65	
Db	235	GKPTINLLASLESCVNLISGHIVELASVIMRSPLEPKFLPDSCILFCSHKYDAVGS	294	
QY	66	-----GMVFYFKCQNNLTQKT	82	
Db	295	YKIQLOMSAQEAGARQMSLSPYSSYSYNDYPPSSLSDIILIRLGNVLNLFYKRYNNYMQET	354	
QY	83	EYREDSCSPFCSMILGSGFKGLQFHLNHSDDLEEEF	118	
Db	355	EYTEDPSCPCFYKCGSGFKQLGCHLANSDDLEFHYEF	390	

```

RESULT 14
US-10-437-963-184120
Sequence 184120, Application US/10437963
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovacic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21453221.8
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 184120
LENGTH: 263
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(263)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_81143C.1.pep
US-10-437-963-184120

```

```

Query Match      8.0%; Score 190; DB 6; Length 263;
Best Local Similarity 53.8%; Pred. No. 2e-09;
Matches 35; Conservative 12; Mismatches 18; Indels 0; Gaps 0;

QY      1 MCRONCRASSPEEYISTDENTLLIYCKPVRILYNIIFHLRISGNSPEFPRCLNTKIKGARRR 60
      ||| :||| : : : :||| ||| ||| :||| ||| ||| ||| :|||
Db      3 MCRHOPRRRLSPDEOLAEESEFALYCKPVELYNIIOIRSRKINPAFLORCLLYIKIHARRK 62

QY      61 KSRST 65
      : | |
Db      63 RSLIT 67

RESULT 15
US-10-424-599-250979
; Sequence 250979, Application US/10424599
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with

```

```

; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 250979
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(195)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MPT3847_68663C.1.pap
US-10-424-599-250979

```

```

Query Match      6.7%; Score 159; DB 6; Length 195;
Best Local Similarity 56.9%; Pred. No. 1.1e-06;
Matches 29; Conservative 10; Mismatches 12; Indels 0; Gaps 0;

QY 11 SPEEYSTDENLLIYCKPVRLYNIFHLRLSLGNPSFLPRLNKKIGAKRRK 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 22 SAEELAAEESLSIYCKPVELYNILQRRMRNPSPFLQRCILHYRIKAKRRK 72

```

Search completed: June 13, 2003, 15:37:12
Job time : 56 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 22:10:23 ; Search time 3586 Seconds

(without alignments)
3611.475 Million cell updates/sec

File: US-09-890-220-2

Sequence: 1 MCRONCAKSPSEVISTDE.....INNNNVNDKNNSDKVIK 445

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: . 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-Q=/cgn2.1/USPTO-spool/US09890220/runat_13062003_144851_16477/app.query.fasta.1.583
-DB=GenEmbl -QFMT=fastap -SUFFIX=trge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45
-DOCCALIGN=200 -THR.SCORE=100 -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220 &cgn2.1.2406 &runat_13062003_144851_16477 -NCPU=6 -ICPU=3
-NO_MAP -LARGEDEV -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_of:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hg_hum:*
31: em_hg_in:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hggo_hum:*
40: em_hggo_mus:*
41: em_hggo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2378	100.0	1722	6	AX032817	AX032817 Sequence
2	2378	100.0	1722	6	AX032888	AX032888 Sequence
3	2370	99.7	1700	8	AF284500	AF284500 Arabidops
4	2363	99.4	1722	6	AX032934	AX032934 Sequence
5	2343.5	98.5	1722	8	AF284501	AF284501 Arabidops
6	2285	96.1	1354	8	AY063047	AY063047 Arabidops
7	2285	96.1	1715	6	AX032818	AX032818 Sequence
8	2285	96.1	1715	6	AX032891	AX032891 Sequence
9	2285	96.1	1721	6	AY034902	AY034902 Arabidops
10	2258.5	95.0	1737	6	AX032894	AX032894 Sequence
11	1813.5	76.3	6338	6	AX032890	AX032890 Sequence
12	1798.5	75.6	91630	8	ATRPSP5LE2	AF180942 Arabidops
13	1659	69.8	5895	6	AX032893	AX032893 Sequence
14	1659	69.8	199548	8	ATCHRIV45	AL161545 Arabidops
15	1659	69.8	201471	8	ATRCAT7	297342 Arabidops
16	923	38.8	2279	8	AB053171	AB053171 Arabidops
17	589	24.8	2284	8	AF096096	AF096096 Arabidops
18	451.5	19.0	558	8	AX032900	AX032900 Sequence
19	400.5	16.8	8508	8	AB053262	AB053262 Arabidops
20	400.5	16.8	87180	8	AB023044	AB023044 Arabidops
21	368.5	15.5	325	6	AX032898	AX032898 Sequence
22	311	13.1	463	6	AX032896	AX032896 Sequence
23	306.5	12.9	135495	2	OS2N01007	AL607099 Oryza sat
24	265	11.1	107436	8	AC006068	AC006068 Arabidops
25	264.5	11.1	6081	8	AF096095	AF096095 Arabidops
26	250	10.5	2040	8	AB053265	AB053265 Arabidops
27	250	10.5	199548	8	ATCHRIV45	AL161545 Arabidops
28	250	10.5	201471	8	ATRCAT7	297342 Arabidops
29	242	10.2	127130	8	OS2N00040	AL606603 Oryza sat
30	240	10.1	113355	8	OS2N00100	AL606654 Oryza sat
31	222	9.3	3346	9	AK074333	AK074333 Homo sapi
32	221	9.3	3942	9	BC015704	BC015704 Homo sapi
33	221	9.3	4441	6	AX032902	AX032902 Sequence
34	221	9.3	4441	6	AX337789	AX337789 Sequence
35	221	9.3	4441	9	D63881	D63881 Human mRNA
36	219	9.2	3648	3	AF149047	AF149047 Drosophill
37	219	9.2	4055	3	AY069809	AY069809 Drosophill
38	208	8.7	83538	2	AC018338	AC018338 Drosophill
39	208	8.7	167816	3	AC009381	AC009381 Drosophill
40	208	8.7	249805	3	AE003515	AE003515 Drosophill
41	178.5	7.5	57873	2	AC010021	AC010021 Drosophill
42	175	7.4	56152	2	AC116363	AC116363 Dictyoste
43	168	7.1	5769	3	AF104350	AF104350 Dictyoste
44	161	6.8	171187	2	AC116960	AC116960 Dictyoste
45	155.5	6.5	12464	2	AC115604	AC115604 Dictyoste

RESULT 1

ALIGNMENTS

AX032817
 LOCUS AX032817 1722 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 57 from Patent WO0046358.
 ACCESSION AX032817
 VERSION AX032817 GI:10279793
 KEYWORDS
 SOURCE
 ORGANISM
 synthetic construct.
 artificial construct
 1 (bases 1 to 1722)
 Dean, C., West, J. and Johanson, U.
 TITLE
 Plant gene
 Patent: WO 0046358-A 57 10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)
 FEATURES
 source
 1. 1722
 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Landsberg erecta VRN2 cDNA"

BASE COUNT 538 a 339 c 348 g 497 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,46e-209 Length: 1722
 Score: 2378.00 Matches: 443
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 100.008 Indels: 0
 Ds: 6 Gaps: 0

US-09-890-220-2 (1-445) x AX032817 (1-1722)

QY 1 Metcysarglnasnancysargalalyserserprogluvalilleserthraspclu 20
 Db 231 ATGTGTAGGCGAATGTCGCGCAATCTCCACGCGAGAGTATTTCAACGATGAG 290
 QY 21 Asnleuuleltyrcyslyspovalarqleutyasnillepheiisleuargserleu 40
 Db 291 AATCTCTGTATATATGTAACTGCTTCACATATATATACATCTTTCACCTTGCGTCTTA 350
 QY 41 Glyasnproserpheleuproargcysleuasnlyrlylleglvalilysarqlysarg 60
 Db 351 GGCACCCATCGTTCTCCACATCTTGAACCTTGAACAAATTTGGAGCAAGCGCAAGA 410
 QY 61 Lysersergethrglymetvalilpheasnlyrlylaspasasnanserthleugln 80
 Db 411 AAGTCAGATCTAGTGGATGCTGATTTCAACTATTAAGATTTGTAATACACANTTACAG 470
 QY 81 Lysrthgluvalarqgluaspcyserserprophesermetleucysgliserphe 100
 Db 471 AAAACGTGAAGTTAGGAGGATTTCTTCTTCATTTTGTCTATCTATGTGTAGCTTC 530
 QY 101 LysgllyleuglnphehisleuasnserSerhisasplleupheglunpheglisleu 120
 Db 531 AAGGCGCTCAATTTCTATTTGATTAATCTCAATTTATTTGATTTGAGTTCAAGCTT 590
 QY 121 Pheglunlutyrglnthrvalasnvalservalylsleuasnserpheillepheiuglu 140
 Db 591 TTGCAAGATACACAGACAGTAAATGTTCTGTAACCTTAATTCCTTATATTTGAGGAA 650
 QY 141 GluylseraspsasplyspheiupropheSerleucysSerlyspoarqlylsarg 160
 Db 651 GAAGGAGTATACATATAATTTGAGCCCTCTCTCTGCTCGAACCCTGTAAGGG 710
 QY 161 Arggluargllyllyragsasnansertharqargleulysvalcyspheleuproleuasp 180
 Db 711 AGACAAAGAGGTGCGAGAAATACACAGAGACTTAAAGTATGCTTTTTCACCGTGGAT 770
 QY 181 SerproserleuthrasnnglythrgluasnglylilethrleuasnaspGlyasnarg 200
 Db 771 TCACCCAGTTTAACTTAATGCGACAGAAATGGAATCACCCCTACTTAATGATGAAACCGT 830

QY 201 Gtyleugllytyrprogluvalatthrululeuallagylgnpheglumethrserasnile 220
 Db 831 GGTATAGATATATCCGAGAGCAACAGACTGCTGGCAAAATTTAGATGACCGCAACATTT 890
 QY 221 ProproalallealeahisSerSerleuaspalaglyalalysvallileuthrserglu 240
 Db 891 CCACAGCCATFAGCCCACTTCTCTGTGAGCGCTGTCTTAATATATATGACCAAGCGAA 950
 QY 241 AlaValalproalathrlysthrArqlylsenserlaqluarqsergluallaargser 260
 Db 951 GCTGTGTCCTGCTACTTAACAGCAAGAAAGTATCTCTGAGGATCAGGAGCGGTGAAAC 1010
 QY 261 HisleuuleuenglnlylserglnpheutyhrhisSerhisrlyalaglnprometlaleu 280
 Db 1011 CACCTACTTCTTCAGAAACGCCCAATTCATCATTTCTCACAGACTCCAGCCANTGGCGTT 1070
 QY 281 GluGlnvalmetSeraspargaspsergluasgluvalaspaspaspvalalasphe 300
 Db 1071 GAGCAAGTATATCTTGACCGGAGATAGCGAGATGATGATGATGATGATGATGATGATG 1130
 QY 301 Gluasparglmetleuaspasphevalaspvalaspvalaspvalaspvalaspvalasp 320
 Db 1131 GAAAGATCGCCAGATCTTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1190
 QY 321 Hisleuutpanserphevalarqlylgnarqvalillealaspclylhislertrp 340
 Db 1191 CATCTTTGACACTGCTTTGTAAAGAAACAAAGGATTATAGCATAGATGATGATGATG 1250
 QY 341 AlaCysglualapheSerargpheyrglylulysgluleuuhisargtyrserleupe 360
 Db 1251 GCATGTGAAGCATTTTCAAGATTTTACAGAAAGATGACCGGTATCATCATCTCTTC 1310
 QY 361 TrpCysrtrparqleupheleulileuGlnasCysargasnserSeraphrthrthr 380
 Db 1311 TGGTGTGGATGATTTGTTTATTAACATATGAGAACCATGATGATGATGATGATGATG 1370
 QY 381 IleasnasnCysasnThrilleuGlnasCysargasnserSeraphrthrthr 400
 Db 1371 ATCAACAACCTGAATACCATCATCTCGAATAATTCGCGTAATGCTCAACACCCACACC 1430
 QY 401 AsnasnasnasnSerValaspargProSeraspserasnansasnasnilleval 420
 Db 1431 AACCAACAACAAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1490
 QY 421 AspHisproasnaspilleasnasnlylaspasnvalaspasnlylaspasnasnserarg 440
 Db 1491 GATCATCCCATGACATTAACAACAACAACAATGTTGACACAAGACATTAACAGCAGA 1550
 QY 441 Asplysvalillelys 445
 Db 1551 GACAAAGTAAATTA 1565

RESULT 2
 LOCUS AX032888 1722 bp DNA linear PAT 22-SEP-2000
 DEFINITION Sequence 1 from Patent WO0044918.
 ACCESSION AX032888
 VERSION AX032888 GI:10279825
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1722)
 Dean, C. and Gendall, A.
 TITLE
 Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 Patent: WO 0044918-A 1 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source
 1. 1722
 Location/Qualifiers

/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 538 a 339 c 348 g 497 t
ORIGIN

Alignment Scores:
Pred. No.: 1,46e-209 Length: 1722
Score: 2378.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-890-220-2 (1-445) x AX032888 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 231 ATGTGTAGGCAAAATTTGTCGCGCAAAATCTCACGAGGAAAGTATTCACACTGATGAG 290
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 291 AATCTCTGATATATTGTAACCTGTGCACTATATACATCTTTCACCTTGCTCTCTTA 350
QY 41 GlysProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 351 GGCAACCCATCGTCTCTTCCAGAGTCTTGAACTACAAAATTGGACAAAGCGCAAAAGA 410
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 411 AAGTAAAGATCTACAGGGATGGATGTTTCAACTAAAGATTTGTAATACACATTTACAG 470
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 471 AAAACTGAAGTTAGGAGAGATGTTCTTCTCCATTTTGTCTATGCTATGTTAGTACTTC 530
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluLysLeu 120
DB 531 AAGGGGCTGCATTTTCATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 590
QY 121 PheGluGluTyrGluThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 591 TTGCAAGAAATACAGACAGTAAATGTTTGTAAACTTAATCTTCAATTTTGGAGAA 650
QY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 651 GAAGAGAGTGAAGAGATTAATTGAGCCCTCTCTCTGCTCGAAACCTCGTAGCGG 710
QY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 711 AGACAAAGAGGTGGCAGAAATTAACACAGAGACACTTAAAGTATGCTTTTACCGTTGAT 770
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
DB 771 TCACCCAGTTTAACATGATGACACAGAAATGGAATCCCTACTATATATGAAACCCGT 830
QY 201 GlyLeuGlyTyrProGluValThrGluLeuAlaGlyGluPheGluMetThrSerAsnIle 220
DB 831 GGTTTAGGATATCCGAGGCAACAGAGCTTGTCTGACATTTGAGATGACGCAACATT 890
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 891 CCACGAGCATAGCCACTCTTCTGAGCGCTGGTGCTAAAGTTATATTGACAAACGAA 950
QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 951 GGTGTGGTCCCTGCTACTAAGACAAAGAAAGTATCTGCGACGAGTACAGAGGCTTAAC 1010
QY 261 HisLeuLeuGluGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1011 CACCTACTCTTCAGAAAGCCCAATTCTATCTACAGAGTCCAGCCAAATGGCGCTT 1070
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300

DB 1071 GAGCAAGTAATGTCTGACCGGATAGCAGAGATGAAGTGCATGACGATGTGCAGATT 1130
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAspLysAspLysGlnPheMet 320
DB 1131 GAAGATCGCCAGATGCTTATGACTTTGTGATGTGAATGAATGAAGATAAGCAATTCAG 1190
QY 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerThr 340
DB 1191 CATCTTGGAACTCGTTTGTAAAGAAACAAAGGTTATGACGATGCTATATCTTGTG 1250
QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGlyLeuHisArgTyrSerSerLeuPhe 360
DB 1251 GCATGGAAGCAATTTTCAAGATTTTACAGAAAGAGTTGACACCGTACTCATCAGCTTTC 1310
QY 361 TrpCysTrpArgLeuPheLeuLeuLysLeuThrAsnHisGlyLeuValAspSerAlaThr 380
DB 1311 TGGTGTGGAGATGTTTGTGAATTAACATATGAAACAGAGCACTGTGCACACGAC 1370
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
DB 1371 ATCAACAATGCAATGCATCTCTCGAATTCGCCATATAGCTCAGACACACACACAC 1430
QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
DB 1431 AACAAACAACAAGATGGATGCTGCCAGTACTCAAAACCAACAACATTAACATTTG 1490
QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB 1491 GATCATCCCAATGACATTAACAACAAGAAATGTTGACAAACAGACATTAACACACA 1550
QY 441 AspLysValIleLys 445
DB 1551 GACAAAGTAATTTAA 1565

RESULT 3
AF284500 1700 bp mRNA linear PLN 21-NOV-2001
LOCUS Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete cds.
DEFINITION
ACCESSION AF284500
VERSION AF284500.1 GI:16945787
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
REFERENCE 1 (bases 1 to 1700)
AUTHORS Gendall, A.R., Levy, Y., Y., Wilson, A. and Dean, C.
TITLE The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis
JOURNAL Cell 107 (4), 525-535 (2001)
MEDLINE 21575875
PUBMED 11719192
REFERENCE 2 (bases 1 to 1700)
AUTHORS Gendall, A.R., Levy, Y., Y. and Dean, C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
FEATURES
source
1..1700
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="IV"
/map="adjacent to RPP5"
/note="ecotype: Landsberg erecta"
1..1700
/gene="VRN2"
227..1564
/gene="VRN2"
/codon_start=1
/product="vernalization 2 protein"
/protein_id="AA132135.1"

```

/db xref="GI:16945788"
/translation="MCRONCRASKEPSEYISIDENLLITCKPVRILYNIFHRISGNS
ELPRCLNFKIGAKRKRKSRSTGMYENKYDCNNITQKTEVEDDSCSPCSNLCSTFG
LOPHINSNDLPPEPKISEEYOTVAVSKINSFIIEEGSDDDPEFSLCSFKR
ROGGRNNRRLKVCZPLIDSPSLNGSLITLNDGNRGLGTEPELGLGOFEMTS
NTPPALNSLDGAKVITLSEAVPATRKLSEASEASHLLQKQFHSRQ
PMALBOVMSRDEDEVDVADFEEDRQDLDFVDNDEQFMHMSFYKQKRIA
DGHISMAECASFSTFKELHRTSLFCWCRFLIKLMHGLVDSATINNCTITLNCR
NSSDTTNNNSVDRPSDNTNNNIVDHPNDINNNKNNVNDKONNSDKYIK"
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="a"
1195
variation
/ gene="VRN2"
/ note="VRN2-1 mutation results in TGG Trp to TGA stop"
/ replace="a"
1399
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="g"
1487
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="a"
1488
variation
/ gene="VRN2"
/ note="compared to the sequence in GenBank Accession
Number AF180942"
/ replace="g"
BASE COUNT      533 a      334 c      347 g      486 t
ORIGIN
Alignment Scores:
Pred. NO.:      7.83e-209      Length:      1700
Score:          2370.00      Matches:      444
Percent Similarity: 99.78%      Conservative: 1
Best Local Similarity: 99.78%      Mismatches: 1
Query Match:     99.66%      Indels:      0
DB:              8      Gaps:        0
US-09-890-220-2 (1-445) x AF284500 (1-1700)
QY      1 MetCysArgGlnAsnGlySerArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
Db      227 ATGTGTAGGAGAGATGTCGCGCAATCCTCACCGAGAAAGATTCATCAATGATGAG 286
QY      21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
Db      287 AATCTCTGATATATGTAAACCTGCTGCATATTAACATCTTCACTTCGCTCTCTA 346
QY      41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
Db      347 GGCACCCATCGTTTCTTCAGATGCTGGAACCTCAAAATTTGAGCAAAACGCAAAAGA 406
QY      61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
Db      407 AACTCAAGATCTACTGGATGTGTAAGTTTCAACTTAAGATTGATTAACACATTTCAG 466
QY      81 LysThrGluValArgGlnAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
Db      467 AAAACTCAAGATGTAGGAGATGTGTCTGTCTCTCATGCTATGCTATGAGGAGCTTC 526
QY      101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
Db      527 AAGGGCTGCATATTCATTTGAATTCATCTCATGATTTGATTTGAACTTCACCTT 586
QY      121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
Db      587 TCGAAGAAATACACAGACGTTAATGTTCTGTAAAAAATTAATTCCTCATATTTGAGGAA 646

```

```

QY      141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
Db      647 GAAGGAGAGGAGAGAGATTAATTTGAGCCCTCTCTCTGCTGCAAACTCGTAAGCGG 706
QY      161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
Db      707 AACACAAAGGAGGAGGAGAAATTAACACAGGAGAGATTAATTAATGCTTTTACCGTGGAT 766
QY      181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
Db      767 TCACCCAGATTAACTAAATGAGCAGAAATGAAATCCCTACTTAATATGAAACCGT 826
QY      201 GlyLeuGlyTyrProGluAlaThrGluIleuAlaGlyGlnPheGluMetThrSerAsnIle 220
Db      827 GGTTAGGATATCCGAGGAGCAACAGAGCTGCTGCAATTTGATGATGACCAACACATT 886
QY      221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
Db      887 CCACGACGATAGGCCACTCTCTCTGAGCGCTGCTCTAAATTAATTTGACCAAGCGAA 946
QY      241 AlaValValProAlaThrLysThrArgLysLeuSerLagLysSerGluAlaArgSer 260
Db      947 GCTGTGTCCTCCGCTACTTAACACAGAAAGTATCTCTGAGCATCAGAGCTAGGCTAGAAC 1006
QY      261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
Db      1007 CACCTACTTCTTCAGAAACGCCCAATTCATCATCTTCACAGAGTCCAGCCAAATGGCGCTT 1066
QY      281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
Db      1067 GAGCAAGTAAATGTCTGACCGGATAGCGAGATGATGATGATGATGATGATGATGATGAT 1126
QY      301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
Db      1127 GAAAGATCCGCAAGATGCTTGAATGACTTGTGATGTGATGAATTAACATGAATTAATC 1186
QY      321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspLysHisIleSerTrp 340
Db      1187 CATCTTTTGAAGCTCGTTTGTAAAGAAACAAAGGGTTATAGCAATGCTATATTTTCGGGG 1246
QY      341 AlaCysGluAlaPheSerArgPheTyrGluLysGluIleuHisArgTyrSerSerLeuPhe 360
Db      1247 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGACTGACCCGCTTACTCATCACTCTTC 1306
QY      361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrAsnHisGlyLeuValAspSerAlaThr 380
Db      1307 TGGTGTGAGATTTGTTTGTATTAATCTATGAAACCATGACCATGACTTGTGACTCACGCC 1366
QY      381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
Db      1367 ATCAACAACTGCATATACATCTCGAATATTCGCGTAATAGCTCAACACCAACCAACCC 1426
QY      401 AsnAsnAsnAsnSerValAspArgProSerArgSerAspThrAsnAsnAsnIleVal 420
Db      1427 AACAAACAACAACATGTGAGATCGTCCAGATCAACCAACCAACAACAATAAATTTGTG 1486
QY      421 AspHisProAsnAspIleAsnAsnLysAsnAspValAspAsnLysAspAsnSerArg 440
Db      1487 GATCATCCCAATCAATTAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAA 1546
QY      441 AspLysValIleLys 445
Db      1547 GACCAAGTAATTAATA 1561
RESULT 4
AX032934
LOCUS      AX032934      1722 bp      DNA      linear      PAT 21-SEP-2000
DEFINITION Sequence 47 from Patent WO0044918.
ACCESSION      AX032934
VERSION      AX032934.1 GI:10279842
KEYWORDS      thale cress.
SOURCE

```


ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Dean, C. and Gendall, A.
TITLE Methods and means for modification of plant characteristics using the vernalization gene vrn2
JOURNAL Patent: WO 0044918-A 47 03-AUG-2000;
DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD (GB)

FEATURES
source Location/Qualifiers
1..1722
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

BASE COUNT 539 a 339 c 347 g 497 t

ORIGIN

Alignment Scores:
Pred. No.: 3.52e-208 Length: 1722
Score: 2363.00 Matches: 444
Percent Similarity: 99.788 Conservative: 0
Best Local Similarity: 99.788 Mismatches: 1
Query Match: 99.37% Indels: 0
DB: Gaps: 0

US-09-890-220-2 (1-445) x AX032934 (1-1722)

QY 1 MetCysArgGlnAsnGlnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 231 AGGTGTAGGCGAGAAATGTCGCGGAAATCTCCACCGAGAGAGTATTCACACTGTGTGAG 290
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 291 AATCTCTATATATTGTAACCTGTTCCACTATATACATCTTTACCTCGCTCTCTA 350
QY 41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 351 GCGAACCCATCGTTCTCCAGATGCTTGAACACTACAAATGCGACCAAGCGCAAAAGA 410
QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 411 AAGTCAAGATCTACTGGGATGTGATGTTTCACTAAGAGATTGTATACACATTTACAG 470
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 471 AAACTGAGATTAGGAGAGATGTTCTTCTGCAATTTGCTCTATGATGTGTACTTTC 530
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 531 AAGGGCTGCAATTTCAATTTGAATCATCATCATGATTTATTTGAATTTGAGTTCAAGCTT 590
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
DB 591 TTCGAAGATFACACAGACACTTAATGTTCTGTAATAACTTAATCTTCAATATTGAGGAA 650
QY 141 GluGluSerAspAspLysPheGluPheGluPheSerLeuGlySerLysProArgLysArg 160
DB 651 GAAGGAGATGATGACATATAATTTTGAAGCCCTTCTCTGCTGCGAAGCCCTGCTAAGCGG 710
QY 161 ArgGlnArgGlyValArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 711 AGCAACAAGAGGTGGCGAATAATACACAGAGAGCTTAAGTATGCTTTTACCGTTGAGT 770
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyTleThrLeuLeuAsnAspGlyAsnArg 200
DB 771 TCACCCAGATTAACTATGCGACAGAGAAATGGAATACCCCTACTTAATGATGGAACCGCT 830
QY 201 GlyLeuGlyTyrProGluLarThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 831 GGTTTAGGATATCCCGAGGCAACAGAGCTTGTCTGGCAATTTTGAGATGACCAACACATT 890
QY 221 ProProAlaIleLeuHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240

DB 891 CCACGAGCATATGCCACACTCTCTCGAGCGGTGGTGTAAAGTATATGACAAAGCGGA 950
QY 241 AAlaValProAlaThrIleTyrThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 951 GGTGTGCTCCCTGCTACTAAGACAGMAAGTATATGCTGACGAGCAAGGCTTAAAC 1010
QY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1011 CACCTACTTCTTCAGAAAGCCCAATTTCTCATTTCTCACAGAGTCCAGCAATGGCGCTT 1070
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
DB 1071 GAGCAAGTATATGCTGACCGGGATAGCGAGATGAAGTGCATGATGTCAGATTTT 1130
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB 1131 GAAGATCGCCAGATGCTTATGATGACTTGTGAGATGTGAATGAATGAATGAATGAAT 1190
QY 321 HisLeuTyrPasnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB 1191 CATCTTTGAAACTCGTTTGTAGAAACAAAGGCTTATAGCATGCTATATCTTGTG 1250
QY 341 AlaCysGluAlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAGAGTTGCACCGTTACTCATCATCTTTC 1310
QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrPasnHisGlyLeuValAspSerAlaTrp 380
DB 1311 TGGTGTGGAGATGTTTTGTGATTAACATATGAAACAAAGGCTTGTCCACGACGAC 1370
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAspSerSerAspThrThrThr 400
DB 1371 ATCAACAACCTGCAATATCCATCCCGAGATTTGCCGTAATAGCTCACACACACACAC 1430
QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerSerAsnThrAsnAsnAsnIleVal 420
DB 1431 AACCAACAACAAGTGTGATGCTGCCAATGACTCAAAACCAACAACATTAACATTTGTG 1490
QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB 1491 GATCATCCCAATGACATTAACAACAAGAAACAAATGTTGACAAACAGGACATTAACAG 1550
QY 441 AspLysValIleLys 445
DB 1551 GACAAAGTATTTAA 1565

RESULT 5
AF284501 1722 bp mRNA linear PLN 21-NOV-2001
LOCUS Arabidopsis thaliana vernalization 2 protein (VRN2) mRNA, complete
DEFINITION cds, alternatively spliced.
ACCESSION AF284501
VERSION AF284501.1 GI:16945789
KEYWORDS
SOURCE Arabidopsis thaliana.
ORGANISM Arabidopsis thaliana.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 1722)
AUTHORS Gendall, A.R., Levy, Y.Y., Wilson, A. and Dean, C.
TITLE The VERNALIZATION 2 Gene Mediates the Epigenetic Regulation of Vernalization in Arabidopsis
JOURNAL Cell 107 (4), 525-535 (2001)
MEDLINE 21575875
PUBMED 11719192
REFERENCE 2 (bases 1 to 1722)
AUTHORS Gendall, A.R., Levy, Y.Y. and Dean, C.
TITLE Direct Submission
JOURNAL Submitted (03-JUL-2000) Molecular Genetics, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK
FEATURES
Location/Qualifiers

source	1. 1722	/organism="Arabidopsis thaliana"
	/db_xref="taxon:3702"	
	/chromosome="1v"	
gene	/map="adjacent to RPPs"	
	/note="scotype: Landsberg erecta"	
CDS	1..1722	
	/gene="VRN2"	
	227..550	
	/gene="VRN2"	
	/note="alternatively spliced"	
	/codon_start=1	
	/product="vernalization 2 protein"	
	/protein_id="AL32136.1"	
	/db_xref="GI:16945790"	
	/translation="MCRONCRRAKSSPEEVISTDENLLIYCKPVLVNIIFILRLGNPS FLPRCLNTYIGAKRRKRSKSTGMVYFNKDCNNTLOKTEVREDCSFCPSMLCSFVY GNYYN"	
BASE COUNT	540 a 338 c 353 g 491 t	
ORIGIN	GNYYN	
Alignment Scores:	2.21e-206	Length: 1722
Pred. No.:	2343.50	Matches: 444
Score:	98.01%	Conservative: 0
Percent Similarity:	98.01%	Mismatches: 1
Best local Similarity:	98.01%	Indels: 8
Query Match:	98.55%	Gaps: 1
DB:	8	
US-09-890-220-2 (1-445) x AF284501 (1-1722)		
OY	1 MetCysArgGlnAsnGlySerProGluValIleSerThrAspGlu	20
DB	227 ATGTGAGGAGAGATGTCGCCGGAATCTCCAGGAGAGATTCACATGATGAG	286
OY	21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu	40
DB	287 AATCTCTGATATATGTTCAACCTGTCGACTATATACATCTTTCACCTTCGCTCTCTA	346
OY	41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg	60
DB	347 GGCAACCCATCGTTCTTCACAGATGCTGAACCTCAAAATTTGGAGCAAAAGCAAAAGA	406
OY	61 LysSerArgSerThrGlyMetValIlePheAsnTyrLysAspCysAsnAsnThrLeuGln	80
DB	407 AAGTCAGATCTACCTGGATGTGATTTCAACTTAAGGATTTGAATTAACACATTTACAG	466
OY	81 LysThrGluValArgLysAspCysSerCysPropheCysSerMetLeuCysGlySerPhe	100
DB	467 AAAACCTGAAGTGAAGGAGATGTTCTTCTTCATCTTCATCTTCATCTTCATCTTC	526
OY	101 Lys-----GlyLeuGlnPheHisLeuAsnSerSerHisAspLe	113
DB	527 AAGGTGGCAACTATTACACTGAGGCGCTGCATTTGATTTGATTCATCTCATGATTT	586
OY	113 upheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe	133
DB	587 ATTGAATTTGAGTTCACCTTCGGAACAATACACAGATTTATGTTCTTGTAAAACT	646
OY	133 uAsnSerPheIlePheGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu	153
DB	647 TAATCTCTTCATATTTGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	706
OY	153 uCysSerLysProArgLysArgLysArgLysArgLysArgLysArgLysArgLysArgLys	173
DB	707 CTGCTCGAAACCTCGTAACCGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAA	766
OY	173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleTh	193
DB	767 AGATGCTTTTACCGTTGATTCACCAAGTTAACTAATGACAGAGAAATGGAAATCAC	826
OY	193 tLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluValThrGluLeuAlaGlyI	213
DB	827 CCTACTTAATGATGGAACCGTGTGTTTAGATATCCCGAGCAACAGAGCTTCTGGACA	886
OY	213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAl	233
DB	887 ATTGAGATGAGACCAACCAATTCACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	946
OY	233 aLysValIleLeuThrSerGluAlaValIleProAlaThrLysThrArgLysLeuSerAl	253
DB	947 TAAAGTTATTTGAGCAACCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1006
OY	253 aGluArgSerGluAlaArgSerHisLeuLeuGluGlnLysArgGlnPheTyrHisSerHis	273
DB	1007 TGAGGAGATGAG	1066
OY	273 sArgValGlnProMetAlaLeuGluGluGluGluGluGluGluGluGluGluGluGluGlu	293
DB	1067 CAGAGTCCAGCAACAG	1126
OY	293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs	313
DB	1127 CGATGACCATGTTGACATTTTGAAGATCCGACATCTTGATGATCTTGTGATGTGAA	1186
OY	313 nLysAspGluLysGlnPheMetHisLeuTyrPasnSerPheValArgLysGlnArgValI	333
DB	1187 TAAAGATGAAAGCAATTCATGATCTTTGGAAGTGGTTGTGAAGAAACAAAGGTTAT	1246
OY	333 eAlaAspGlyHisIleSerThrPalaCysGluAlaPheSerArgPheTyrGluLysGluLe	353
DB	1247 AGCAGATGCTCATATTTGCTGTCGATGAGATGAGATTTTCAAGATTTTACGAAAGAGTT	1306
OY	353 uHisArgTyrSerSerLeuPheTyrPCTTPATGLeuPheLeuIleLysLeuTyrPasnH	373
DB	1307 GCACCGTACTCATACCTCTCTGCTGTGTGAGAGAGTTTGTATTAACATGAGAACCA	1366
OY	373 sGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAs	393
DB	1367 TGGACTTTCAGATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1426
OY	393 nSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAs	413
DB	1427 TAGCTCAGAC	1486
OY	413 nThrAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAs	433
DB	1487 CACCAACCAACATTAACATTTGATGATCCCAATGACATTAACACACACACACACAT	1546
OY	433 pAsnLysAspAsnAsnSerArgAspLysValIleLys 445	
DB	1547 CAACAGAGACAAATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	1583
RESULT 6		
AY063047	1354 bp	mrna, 11near, PLN 18-APR-2002
LOCUS	Arabidopsis thaliana unknown protein (Atg16845) mrna, complete	
DEFINITION	cds.	
ACCESSION	AY063047	1 GI:17104664
VERSION	AY063047.1	
KEYWORDS	FLI CDNA.	
SOURCE	Arabidopsis thaliana.	
ORGANISM	Arabidopsis thaliana.	
REFERENCE	Arabidopsis thaliana.	
AUTHORS	Arabidopsis thaliana.	
TITLE	Arabidopsis thaliana.	
JOURNAL	Arabidopsis thaliana.	
REFERENCE	Arabidopsis thaliana.	
AUTHORS	Arabidopsis thaliana.	

TITLE
JOURNAL

COMMENT

Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shin, P., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
Direct Submission
Submitted (15-NOV-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

FEATURES

SOURCE

The Salk, Stanford, GSEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Yamada, K., Banb, J., Banno, F., Chang, E., Dale, J.M., Goldsmith, A.D., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wu, H.C., Yamamura, Y., Yu, G., Yu, S., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koesema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shin, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/GSEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/GSEC) contributed equally to this work as PIs.

Location/Qualifiers

1..1354
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="U09871"
/note="This clone is in pUNI 51.
ecotype: Columbia"
1..1354
/gene="At4g16845"
1..1323
/gene="At4g16845"
/codon_start=1
/evidence="experimental
/product="unknown protein"
/protein_id="AAL34221.1"
/db_xref="GI:17104665"
/translation="MCRONCRKSSPEYISTDENLLYCKRYLNTFHLRLSNPS
FLPCLNTKIGAKRRKRSRTGMVFNKDCNNTLQRTVEKDDSCSPCSMLCGSFKG
LQFLNLSHDLFEFEKFLLEBYQTVNVAVKNSLFEEDSGDDDEFEPSLCSKPKR
RORGRNNTRLKVCFLPLDPSLNGTENGIALINDNGRIGYBEAEELAGOFMTS
NIPPAIASSLDAGAKVITTEAVVPAATKRLTSEKSESHLLIORKOFHSHRVO
PMALBOVNSDDSDDEVDVDAFEDROMLDFVDVNDKQFPHILMSVPKQKVIA
DGHISMACEVSRFEKELHCTSLFKCMRLFLIKLHNGLDVSDATTINCNITLSENR
NTSVTNNNNSVDHPSDSNTNNNNIIVDHPNDIKNNVNDKNNSDRK"

CDS

gene

3' UTR

BASE COUNT 431 a 264 c 287 g 372 t
ORIGIN

Alignment Scores:

Pred. No.: 3.94e-201 Length: 1354
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 96.09% Indels: 2
DB: 8 Gaps: 1

US-09-890-220-2 (1-445) x AY063047 (1-1354)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 1 ATGTGTAGGCGAATTTGTCGCGGAAATCTCACCGAGGAGTATTTCAACTATGAG 60
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40

|||||
61 AATCTCTGATATATGTAACCGTCGTGACATATTAACATCTTCACCTCGCTCTTA 120
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 121 GGCAACCCATGCTGTTCTGCCAAGATGCTTGAACCTCAAAATTTGGGGCAAGCGCAAGA 180
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 181 AAGTAAAGATCTACTAGGATGGTATGTTTCAACTTAAGATGTAATTAATACATTACAA 240
QY 81 LysThrGluValArgLysAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 241 AGAAGTGAAGTATAGGAGAGATGTTCTTCTGTCATTTCTCTATTCATGATGATGCTTC 300
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGluLysLeu 120
DB 301 AAGGGCGTCGCAATTTCAATTTGAATTCATCTCATGATTTGAATTTGAGTTCAAGCTT 360
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 361 TTGGAAGATACAGACAGATTAATGTTCTGTAAACCTTAATTCCTTCATATTGAGGA 420
QY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 421 GAAGAGATGATGATGATTAATTTGAGCCCTTCTCTCTGCTGCAAACTGTAAGCGT 480
QY 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 481 AGACAAGAGGCGGAGAAATGAACACAGGAACTTAAGTATGCTTTTACCGTTGGAT 540
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
DB 541 TCACCAAGTTTATGCTTAATGACACAGAAATGGAATGGCTTCCTGAATGATGAAACCGT 600
QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 601 GGTATGAGTATCCGAGGACACAGAGCGTCTGTCGCAATTTGAGATGACTGCAACAT 660
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 661 CCACGAGCATGCGCCACTCTCTGACGCTGTGAAAGTATATTAACAACCGAA 720
QY 241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 721 GCTGTGTCCTGCTACTAAGACAAAGAAATTAATGCTGACCGATCAGAGGCTAGAAC 780
QY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 781 CACCTACTCTTCAGAAACGCCAATTCATATCTTCACAGAGTCCAGCAATGGCGCT 840
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
DB 841 GAGCAAGTAATCTGATCGAGTACGCGAGTGAAGTCGAGAGAGATGTTCCAAATTTT 900
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB 901 GAAAGATCGCCATGCTGATGACTTTGTGATGTGAATTAAGATGAAGAAATTCATG 960
QY 321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 340
DB 961 CATCTTTGGAACCTGTTGTAAGAAACAAAGGTTATAGAGATGTCATATCTCTGG 1020
QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1021 GCATGTGAAGTATTTTCAAGATTTTACGGAAGAGTTGCACTGTACTCATCACTCTTC 1080
QY 361 TrpCysTyrArgLeuPheLeuIleLysLeuTyrPheAsnHisGlyLeuValAspSerAlaThr 380
DB 1081 TGGTGTGAGATTTTATTTGATTAACATGAAACCATGCACTGTGACTCAACCCACC 1140
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
|||||

Db 1141 ATGACACCTGCAATCCATCTCCGAGAAATTCGCTAATACCTCA-----GTCACTAAC 1194
 QY 401 AAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 Db 1195 AACACACACACAGCTGTGATCCCTGATCTCAACACACACACAAATACATTTGTTG 1254
 QY 421 AspHisProAsnAspIleAsnAsnIysAsnAsnValAspAsnIysAspAsnSerArg 440
 Db 1255 GATCATCCGATGATACATTAATAAACAAGAAATGTTGACAAACAGACAAATACACAGCA 1314
 QY 441 AspIys 442
 Db 1315 GACCAAG 1320

RESULT 7
 AX032818
 LOCUS AX032818 1715 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 58 from Patent WO0046358.
 ACCESSION AX032818
 VERSION AX032818.1 GI:10279794
 KEYWORDS
 SOURCE synthetic construct.
 ORGANISM
 REFERENCE
 AUTHORS Dean,C., West,J. and Johanson,U.
 TITLE Plant gene
 JOURNAL Patent: WO 0046358-A 58 10-AUG-2000;
 DEAN CAROLINE (GB) ; WEST JOANNE (GB) ; PLANT BIOSCIENCE LTD (GB) ;
 JOHANSON URBAN (SE)

FEATURES
 source
 1. 1715
 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Columbia VRN2 cDNA"
 BASE COUNT 536 a 326 c 353 g 500 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.42e-201 Length: 1715
 Score: 2285.00 Matches: 428
 Percent Similarity: 97.518 Conservative: 3
 Best Local Similarity: 96.838 Mismatches: 9
 Query Match: 6 Indels: 2
 Gaps: 1

US-09-890-220-2 (1-445) x AX032818 (1-1715)

QY 1 MetCysArgGlnAsnCysArgAlaIysSerSerProGluGluValIleSerThrAspGlu 20
 Db 233 ATGTGTAGCGCAGAAATTTGCGCGCAATCTCACCGAGAGATTCACATGATGAG 292
 QY 21 AsnIleuLeuIleTyrCysIysProValArgLeuTyrAsnIlePheHisIleAspGlySerLeu 40
 Db 293 AATCTGTGATATGATGAAACCTGTGACATAATAACAATCTTCACCTTCGCTCTCTA 352
 QY 41 GlyAsnProSerPheLeuProArgCysIleuAsnTyrIysIleGlyAlaIysArgIysArg 60
 Db 353 GCGAACCCATCTCTTCTCGCAAGAGCTTGACACACAAATTTGGCGCAAGAGCAAGAGA 412
 QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrIysAspCysAsnAsnThrLeuGln 80
 Db 413 AAGTCAGATCTTACTGGGTGGTGTACTTTTCAACTATAGAGATTGTATTAATACATTACAA 472
 QY 81 LysThrGluValAlaGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
 Db 473 AGAAGCTGAAGTAGGAGAGATTGTTCTTCATTTTGTCTATGCTATGTGGTAGCTTC 532
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheIysLeu 120
 Db 533 AAGGGGCTGCAATTCATTTCAATTCATCATCATCATATTTATTTGAATTTGAGTTCAAGCTT 592
 QY 121 PheGluGluTyrGlnThrValAsnValSerValIysIleuAsnSerPheIlePheGluGlu 140

Db 593 TTGGAGAGATACCCAGACAGTTAAATGTTGTGAATACTTAATCTTCATTTTACAGGA 652
 QY 141 GluIlySerAspAspAspIysPheGluProPheSerLeuCysSerIysProArgIysArg 160
 Db 653 GAAGGAATGATGATGATAAATTTTGACCCCTTCTCTCTGCTCGAAACCTGTAGAGCT 712
 QY 161 ArgGlnArgGlyIysArgAsnAsnThrArgArgLeuIysValCysPheLeuProLeuAsp 180
 Db 713 AGCAAAAGAGGTGGCAGAAATAAACACGAGAGACTTAAGTATGCTTTTACCGTTGAT 772
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspIlysAsnArg 200
 Db 773 TCACCCAGTTTAGCTAATGGCAGAAATGGAATGGCCCTGCTGAATGATGAAACCGT 832
 QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
 Db 833 GGTTAGAGATATCCCGAGGCAACAGACTTGTGACAAATTTGATGACTAGCAACAT 892
 QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaIysValIleLeuThrSerGlu 240
 Db 893 CCACAGCCATAGCCCACTCTTCTGTGACGCTGGTGCTAAAGTTATATTACACCGGA 952
 QY 241 AlaValAlaProAlaThrIysThrArgIysLeuSerAlaGluAsnSerGluAlaArgSer 260
 Db 953 GCTGTGCTCCCTGCTACTAAGACAGAAAGTTATCTGTCGACGATCAGAGCTGAAAGC 1012
 QY 261 HisLeuLeuLeuGlnIysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 Db 1013 CACCTACTTCTTCAAGAAAGCCCATCTCATCTTCACAGAGTCCAGAGCCAAATGGCCCT 1072
 QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
 Db 1073 GAGCAAGTAATGTGTGATCGGATGAGAGATGAAAGTGCATGATGCTGCTGACATTTT 1132
 QY 301 GluAspArgIleMetLeuAspAspPheValAspValAsnIysAspGlnIysGlnPheMet 320
 Db 1133 GAAGATCCCGCAGATGCTGATGACTTGTGATGATGAATAAAGATGAAAGCAATTCATG 1192
 QY 321 HisLeuThrPAsnSerPheValArgIysGlnArgValIleAlaAspGlyHisIleSerTrp 340
 Db 1193 CATCTTTGGAACCTGTTGTGTAAGAAACAAAGGCTTATGACAGATGCTCATCTCTGG 1252
 QY 341 AlaCysGluAlaPheSerArgPheTyrGlyIysGlyLeuHisArgTyrSerSerLeuPhe 360
 Db 1253 GCATGTGAAGTATTTCAAGATTTTACGAGAAAGAGTTGCACTGTACATCAGACTCTTC 1312
 QY 361 TrpCysTrpArgLeuPheLeuIleIysLeuThrPAsnHisGlyLeuValAspSerAlaThr 380
 Db 1313 TGGTGTGGAGATTTGTTTGAATTAACATATGGAACCATGAGACTTGTGACACCGCAC 1372
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspPheThrThrThr 400
 Db 1373 ATCAACAACCTGCATACCATCTCGAGAAATGCGCTTAATACCTCA-----GTCACTAAC 1426
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleVal 420
 Db 1427 AACACACACACACAGTGGATGATCCCAAGGATCTAAACACCAACACAAATACATTTGTG 1486
 QY 421 AspHisProAsnAspIleAsnAsnIysAsnAsnValAspAsnIysAspAsnAsnSerArg 440
 Db 1487 GATCATCCGAATGACATTAATAAACAAGAAACAAATGTTGACACACAGCAATACACAGCA 1546
 QY 441 AspIys 442
 Db 1547 GACCAAG 1552

RESULT 8
 AX032891
 LOCUS AX032891 1715 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 4 from Patent WO0044918.
 ACCESSION AX032891
 VERSION AX032891.1 GI:10279827

KEYWORDS

SOURCE

thale cress.
Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

ORGANISM

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE

1 (bases 1 to 1715).

AUTHORS

Dean, C. and Gendall, A.

TITLE

Methods and means for modification of plant characteristics using
the vernalization gene vrn2

JOURNAL

Patent: WO 0044918-A 03-AUG-2000;
DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
(GB)

FEATURES

Location/Qualifiers

BASE COUNT

536 a 326 c 353 g 500 t

ORIGIN

/db_xref="taxon:3702"

Alignment Scores:

Pred. No.:	5,426-201	Length:	1715
Score:	2285.00	Matches:	428
Percent Similarity:	97.51%	Conservative:	3
Best Local Similarity:	96.83%	Mismatches:	9
Query Match:	96.09%	Indels:	2
DB:	6	Gaps:	1

US-09-890-220-2 (1-445) x AX032891 (1-1715)

QY 1 MetCysArgGlnAsnGlyAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 233 ATGTGTAGGCAAGATGTCGCGCGAATCCTCACCAGGAGATATTTCACACTATGAG 292
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 293 AATCTCTGATATATTTGTAACCTCTGACATATTAACATCTTTCACCTTGCCTCTCA 352
 QY 41 GlyAsnProSerPheLeuProArgSerLysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 333 GGCACCCATGCTTCTGCGCAAGATGCTTGAACTCAAAATTTGGGCAAGCGCAAAAGA 412
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 DB 413 AAGTCAAGATCTACTGCGGATGCTGTTTCAACCTTAAGATTAATATATCATTAACA 472
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 473 AGAAGTGAAGTTAGGAGGATGTTCTTGTCCATTTTGTCTATATGTTAGTTC 532
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 DB 533 AAGGGCTGCAATTCATTTGATTCATCTCATGATTTATTTGAAATTTGAGTTCAAGCTT 592
 QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 593 TTGGAAGATACCAACAGATTAATGTTCTGTAACCTTAATTCCTTCATATTAGAGA 652
 QY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 DB 653 GAAAGAAAGTGAATGAATTAATTTGAGCCCTTCTCTCTGCTCAAAACCTTGTAAGCGT 712
 QY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
 DB 713 AGACAAAGAGGTGGGAAATTAACACCGAGAGACTTAAGATTAATGCTTTTACCGTTGAT 772
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
 DB 773 TCACCCAGTTTATGATTAATGCGACAGAAATGGAATTTGCCCTGCTCAATGATGAACCGT 832
 QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
 DB 833 GCTTTAGATATATCCGAGGACAGAGAGCTTGTGACAAATTTGAGATGACTGACAAATTT 892

QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 DB 893 CCACAGCCATAGCCACCTCTTCTTGAGCGGTGTCTTAAGATTAATTAACAACCGA 952
 QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 DB 953 GCTGTGTCCCTGCTACTAAGACAAAGAAATTTATGCTGACGATCAGAGCTTGAAAC 1012
 QY 261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 DB 1013 CACTACTTCTTCAAGAAAGCCATTTATCTATCTCTCAGAGTCCACGACATGCGCTT 1072
 QY 281 GluGlnValMetSerAspArgSerGluAspGluValAspAspValAlaAspPhe 300
 DB 1073 GAGCAAGTAATGTCTGATCGGATGCGAGGATGAAGTCGATGAGATGTCACATTTT 1132
 QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 DB 1133 GAAAGATCGCCAGATCTTGATGATCTTTGTGATGATGAATTAAGATGAAGCAATTCATG 1192
 QY 321 HisLeuTyrPasnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
 DB 1193 CATCTTTGGAAGCTCTTGTGTAAAGAAACAAAGGTTATAGCAGATGCTATCTCTTGG 1252
 QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
 DB 1253 GCATGTGAAGTATTTTCAAGATTTTACGAAAGAGTGTGCTGTACTCATCACTCTTC 1312
 QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrPasnHisGlyLeuValAspSerAlaThr 380
 DB 1313 TGTGTGTGAGATGTTGTTTGTATTAACATGAACCATGAGCTTGTGACTCAGCACACC 1372
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 DB 1373 ATCAACAACATGCAATACATCTCGAGATTTCCCTGAATATCACTCA-----GTCACTAAC 1426
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 DB 1427 AACCAACACACAGGTGTGATATCCAGTGTCTCAACACCAACACACATTAACATTTGTG 1486
 QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
 DB 1487 GATCATCGGAATGACATATAAANAAGAAACAAATGTTGACAAACAGACATTAACAGCAGA 1546
 QY 441 AspLys 442
 DB 1547 GACAAAG 1552

RESULT 9
 AY034902 1721 bp mRNA linear PLN 23-APR-2002
 LOCUS AY034902.1 GI:14334421
 DEFINITION Arabidopsis thaliana unknown protein (At4g16845) mRNA, complete cds.
 ACCESSION AY034902
 VERSION AY034902.1 GI:14334421
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana.
 ORGANISM Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 1721)
 Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K.,
 Goldsmith, A.D., Lee, J.M., Quach, H.L., Tortum, M., Yu, G., Bowser, L.,
 Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J.,
 Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, D., Kim, C., Lam, B.,
 Lin, J., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T.,
 Satou, M., Seki, M., Shim, P., Southwick, A., Shinzaki, K.,
 Davis, R.W., Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 TITLE JOURNAL
 REFERENCE 2 (bases 1 to 1721)

AUTHORS

Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shinozaki, A., Southwick, A., Shinozaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.

Direct Submission
Submitted (10-MAY-2001) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN Arabidopsis Full Length cDNA": Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinozaki, K.

TITLE
JOURNAL
COMMENT

The Sak, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Liu, S.X., Sakano, H., Pham, P.K., Banh, J., Chung, M.K., Goldsmith, A.D., Lee, J.M., Quach, H.L., Tang, C.C., Toriumi, M., Yu, G., Bowser, L., Chen, H., Cheuk, R., Jones, T., Karlin-Neumann, G., Kim, C., Koeseema, E., Lam, B., Lin, J., Meyers, M.C., Miranda, M., Nguyen, M., Palm, C.J., Shino, P., Southwick, A., Davis, R.W., Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

FEATURES

source

1. 1721
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="4"
/clone="RAF105-08-K19 (R09811)"
/note="this clone is in a modified phluescript vector (lambda ZAP) as a XhoI/SerI insert."
ecotype: Columbia

gene
5'UTR
CDS

1. 231
/gene="At4g16845"
232. 1534
/gene="At4g16845"
/codon_start=1
/evidence="experimental"
/product="unknown protein"
/protein_id="AAK59409.1"
/db_xref="GI:14334422"
/translation="MCRQNCRAKSSPEVISTDENLTYCKPYRLXNIETRLSAGNS
FLPRLNLTGKRRKRSSTGMYVFNKDCNNLTQRTVEEDSCSPFSLGSGFKG
LQFHLNNSHDLFEFEFLLEBYOTVANSVKLIFEEDSDDKPEFSLGSGPKR
RQGRNNTRRLKVCFLDEBPSLANGENGLANGNGLGYEPATELAGOFEMTS
NIPPAIASSLDAGAKVLTTEAVAPATKRLTSAESESARSHLLQKROFYSEHTA
PALQVMSDRDSEDDVDVADEFEDROMLDDPEVDKDEKORPMHLNVSFRQRTA
DCHISMACEVFSREYKELHCYSLEFCWRLFLIKLNHGLVYSATINNCNTILENCR
NTSVNNNNNSYSDHPSDSENTNNNNIYVHPNDIKKKNVADKONNSRDK"
1555. 1721
/gene="At4g16845"

3'UTR
BASE COUNT
ORIGIN

548 a 324 c 353 g 496 t
Alignment Scores:
Pred. No.: 5 45e-201 Length: 1721
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 96.09% Indels: 2
DB: 8 Gaps: 1

US-09-890-220-2 (1-445) x AY034902 (1-1721)

QY 1 MetCysArgGlnAsnGlySerSerProGluValIleSerThrAspGlu 20
|||||

Db 232 ATGTGTAGGACGAGATTGTCGGCGGCAATCTCCACCGGAGAGACTGATTCATGATGAG 291
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
|||||
Db 292 AATCTTGATATATATGTAACCGTTCGATATATACATCTTCACTCGCTCTCA 351
QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
352 GGCAACCCATCGCTTCTCCCAAGATGCTTGAAACATCAAAATTTGGGCAACCGCAAAATA 411
Db 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
412 AAGTCAGATCTACTGCGGATGATGTTTCACTATAGATAGATGATATATATACATACANA 471
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
472 AGACTGAAAGTTAGGAGAGATTTCTTCTGCAATTTTGTCTATGCTATGCTGATGCTTC 531
Db 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGlyLeu 120
532 AAGGCGCTGCATTTTCATTTGATATTCATTCATGATATTTATTTGATTTGATTTCAAGCTT 591
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
592 TTGGAAGATACCGACAGCTTAATGTTTCTGTAAACTTATTCCTTCATATTGAGANA 651
Db 141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
652 CAGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 711
QY 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
712 AGACAAAGATGTCGACAGAAATTAACACGAGAGCTTAAGTATGCTTTTACCGTGGAT 771
Db 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysArg 200
772 TCACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
832 GGTATTAGATATCCCGAGGACACAGACTTGTGAGCAATTTGAGATGCTGACAACTT 891
Db 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
892 CCACGAGCATAGCCCACTCTCTCGAGCGTGTGTAAGTATATATTAACAACCGAA 951
QY 241 AlaValAlaProAlaThrIleThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
952 GCTGGGCTCCTGCTACTAGACAGAAAGTATATCTGTGACGATCAGAGCTGAGAAC 1011
Db 261 HisLeuLeuGlnLysArgLysArgLysArgLysArgLysArgLysArgLysArgLysArg 280
1012 CACTACTCTTCTGGAAGCCCAATTCATTCATTCACAGATCCAGCAATGCGCTT 1071
QY 281 GluIleValMetSerAspArgAspSerGluAsnGlyValAspAspAspValAlaAspPhe 300
1072 GAGCAAGTAAATGCTGATCGCGATAGCAGAGATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1131
Db 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
1132 GAAATGCGCAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
QY 321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 340
1192 CATCTTTGGAACCTGTTTGAACAAACAAAGGTTTATAGCAGATGTCATATCTCTGG 1251
Db 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
1252 GCATGTAAGTATTTTCAAGATTTTCAAGAAAGAGTTGCTGTTACTTCACTGCTCTTC 1311
QY 361 TyrCysTyrArgLeuPheLeuIleLysLeuTyrAsnHisGlyLeuValAspSerAlaThr 380
1312 TGGTGTGAGATTTGTTTGAATTAACATATGAAACATATGACCTTGTGACTGACCAACC 1371

QY 381 ileasnAsnCysAsnThrIleleuGluAsnCysArgAsnSerSerAspThrThrThr 400
 DB 1372 ATCAACACACTGCAATACCATCTCCAGAAATTCCTGTAATCTCA-----GTCACTAAC 1425
 QY 401 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 DB 1426 AACCAACACACAGTGTGATCATCCAGTGACTCAACACACCAACAACTAATGTTG 1485
 QY 421 AspHisProAsnAspIleAsnAsnIlyAsnAsnValAspAsnIlyAspAsnSerArg 440
 DB 1486 GATCATCCGAATGAGATATAAAACAAAGAACATGTTGACACACAGACATTAACAGACA 1545
 QY 441 AspIlys 442
 DB 1546 GACAAAG 1551
 RESULT 10
 AX032894 1737 bp DNA linear PAT 21-SEP-2000
 LOCUS Sequence 7 from Patent WO0044918.
 DEFINITION AX032894
 ACCESSION AX032894
 VERSION AX032894.1 GI:10279829
 KEYWORDS
 SOURCE
 ORGANISM
 thale cress.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases 1 to 1737)
 Dean,C. and Gendall,A.
 Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 Patent: WO 0044918 A 7 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source 1..1737
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 543 a 330 c 359 g 505 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,53e-198 Length: 1737
 Score: 2258.50 Matches: 428
 Percent Similarity: 95.78% Conservative: 3
 Best Local Similarity: 95.11% Mismatches: 9
 Query Match: 94.97% Indels: 10
 DB: 6 Gaps: 2
 US-09-890-220-2 (1-445) x AX032894 (1-1737)
 QY 1 MetCysArgGluAsnCysArgAlaIlySerSerProGluGluValIleSerThrAspGlu 20
 DB 233 ATGTAGAGCAGAAATTTGCGCGCAATCTCACCGAGAGAGTATTCAACTGATGAG 292
 QY 21 AsnIleuIleIleIyCysIlyProValArgLeuIyTrAsnIlePheHisIleuArgSerIleu 40
 DB 293 AATCTCTGATATATGTGAACCTGTCACATATATACATCTTCACCTTGCTCTCTCA 352
 QY 41 GlysAsnProSerPheIleuProArgCysIleuAsnIyIlyIleGlyAlaIlyArgIlyAsp 60
 DB 353 GGCACCCATCTCTTCTGCCAGAGATGCTTGAACCTACAAATTTGGGGCAAGCGCAAAAGA 412
 QY 61 IySerArgSerThrGlyMetValValPheAsnIyIlyAspCysAsnAsnThrIleuGln 80
 DB 413 AAGTCAAGATCTACGGGATGTAGTTCACACTAATAGAGATTGTAATATACATTACA 472
 QY 81 IyThrIleuValArgGluAspCysSerCysProPheCysSerMetIleuCysIlySerPhe 100
 DB 473 AGAAGTGAAGTGAAGGAGATGTTCTTCCATTTTGTCTATGCTATGTTGTTGCTTC 532

QY 101 Iys-----GlyLeuGlnPheHisIleuAsnSerSerHisAspIle 113
 DB 533 AAGGTGGCACTATTACACTGAGGGCGCTGCAATTTCTTTGATTCATCTCATGATTT 592
 QY 113 nPheGluPheGluPheIlyLeuPheGluGluIyTrGlnThrValAsnValSerValIlySle 133
 DB 593 ATTGAATTTGAGTTCAGAGCTTTTGAGAGAAATACAGACAGTGAATGTTCTGTAAACT 652
 QY 133 uAsnSerPheIlePheGluGluGluIySerAspAspIyPheGluIyProPheSerIle 153
 DB 653 TAAATCCCTTCATATTGTAGAGGAAGGAAGATGATGATTAATTTGAGCCCTTCCTCT 712
 QY 153 uCysSerIyProArgIlySerArgGluArgGlyIlyArgAsnAsnThrArgArgLeuIy 173
 DB 713 CTGCTCGAAACCTCGTAAACCTAGACAAAGAGGTGCGCAAAATTAACACGAGACTTAA 772
 QY 173 sValCysPheIleuProIleuAspSerProSerIleuThrAsnGlyIyThrGluAsnGlyIle 193
 DB 773 AGTATGCTTTTACCGTTGAGATTACCCAGTTTACTATGACACAGAAATGGAATTGC 832
 QY 193 rIleuAsnAspGlyAsnArgGlyLeuGlyIyTrProGluIleThrGluIleuAlaGlyCl 213
 DB 833 CTGCTGAATGATGGAACCGTGTGTAGGATATCCCGAGCAACAGAGCTTGCTGACA 892
 QY 213 nPheGluMetThrSerAsnIleProProAlaIleAlaHisSerSerIleuAspAlaGlyAl 233
 DB 893 ATTGAGATGACTACCAACATTCACACAGCCATAGCCACTTCTCTTGAGAGCTGTGTC 952
 QY 233 alyValIleIleuThrSerGluAlaValAlProAlaThrIyThrArgIlyLeuSerAl 253
 DB 953 TAAAGTTATATTAACAACAGCAAGCTGTGCTCTGCTACTAGACAAAGATTATCTGC 1012
 QY 253 aGluArgSerGluAlaIyArgSerHisIleuIleuGlnIyAspGlnPheIyHisSerI 273
 DB 1013 TTAGGATCAAGAGGTAGAGAGCCACTTCTTTCAGAAACCCCAATTCATCATCTCA 1072
 QY 273 sarGValGlnProMetAlaIleuGluGlnAlaMetSerAspArgAspSerGluAspGluVal 293
 DB 1073 CAGAGTCCAGCCCAATGCGCCTTGACCAAGTATATGCTGATCGGATAGAGAGATGAAGT 1132
 QY 293 lAspAspAspValAlaAspPheGluAspArgGlnMetIleuAspAspPheValAspValAs 313
 DB 1133 CGATGACGATGTTGCGAGATTTTGAAGATCGCCAGATGCTGTGATGACTTGTGATGTGAA 1192
 QY 313 nIlyAspGlyIyGlnPheMetHisIleuTrpAsnSerPheValArgIlyGluArgValIl 333
 DB 1193 TAAAGATGAAAGCAATTCATGCTTTGGAAGTCGTTTGTAAAGAAACAAAGGGTAT 1252
 QY 333 eAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSerArgPheIyGlyIyGluGlyIle 353
 DB 1253 AGCAGATGCTCATATCTCTTGCGCATGTGAAGTATTTTCAGATTTTTCAGAAAGAGTT 1312
 QY 353 uHisArgIySerSerIleuPheTrpCysTrpArgIleuPheIleuIlyLeuTrpAsnIl 373
 DB 1313 GCACGTACTACATACATCTTCTGTTGTGAGATGTTTGTGATTAACATATGAGAACCA 1372
 QY 373 sGlyIleuValAspSerAlaThrIleAsnAsnCysAsnThrIleIleuGluAsnCysArgAs 393
 DB 1373 TGGACTTGTGCACTAGCCACCATCAACAACATGATACATCTCGAGAAATGCGGTAA 1432
 QY 393 nSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAs 413
 DB 1433 TACCTCA-----GTCACTAAACAACAACAACAGTGTGATCATCCAGTACACTCAA 1486
 QY 413 nThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnIlyAsnAsnValAs 433
 DB 1487 CACCAACAACATTAACATGTGATCATCCGATGACATTAATAAACAACAAGATGTGA 1546
 QY 433 pasIyAspAsnAsnSerArgAspIlys 442
 DB 1547 CAACAAGACAAATTAACAGCAGAGACAAG 1574
 RESULT 11

AX032890	AX032890	6338 bp	DNA	linear	PAT 21-SEP-2000
LOCUS	Sequence 3 from Patent WO0044918.				
DEFINITION	AX032890				
ACCESSION	AX032890.1	GI:10279826			
VERSION					
KEYWORDS					
SOURCE	thale cress.				
ORGANISM	Arabidopsis thaliana				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 6338)				
AUTHORS	Dean, C. and Gendall, A.				
TITLE	Methods and means for modification of plant characteristics using the vernalization gene vrn2				
JOURNAL	Patent: WO 0044918-A 3 03-AUG-2000.				
	DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD (GB)				
FEATURES	Location/Qualifiers				
SOURCE	1..6338				
	/organism="Arabidopsis thaliana"				
	/db_xref="taxon:3702"				
BASE COUNT	1932 a 1076 c 1134 g 2193 t	3	others		
ORIGIN					
Alignment Scores:					
Pred. No.:	9,1e-157	Length:	6338		
Score:	1813.50	Matches:	440		
Percent Similarity:	43.58%	Conservative:	1		
Best Local Similarity:	43.48%	Mismatches:	4		
Query Match:	76.26%	Indels:	569		
	6	Gaps:	13		
US-09-890-220-2 (1-445) x AX032890 (1-6338)					
QY	1 MetCysArgLInaSnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20				
DB	2342 ATGTGTGGCAGAAATGTCGCGGAATCTCACCGGAGGAGATTTTCAACGATGAG 2401				
QY	21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40				
DB	2402 AATCTCTGTATATATGTAAACCTGTTCGACTATATAACATCTTCACCTTCGCTCTA 2461				
QY	41 GlyAsn-----				
DB	2462 GGCACGCTGATTTGCTTCCTCTCATCATATTAGCTCAGTAATCTTCANCTCTG 2521				
QY	42 -----				
DB	2522 TGTAGATCACCCACCAATAAGTTGATGTTGCTAAGCTGATTATGCTGATTCAATGGCGA 2581				
QY	43 -----				
DB	2582 GTGTGTCCTCTTTGTCTCTCTAAATTTGAACCTGTGTGTTGTTGTTGTCAGCCATCGTTT 2641				
QY	46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysSer-----				
DB	2642 CTTCACAGATGCTTGAACAAATAATGAGCAAAAGCCAAAGAAAGTA-TGTTTCTT 2700				
QY	62 -----				
DB	2701 CTGGAATGATGCTGCTACAGTATATGTTATTACTTAAATATGAGACGTGATG 2760				
QY	62 -----				
DB	2761 ACCTATTTATCTTGTGTGAGATATGACATAATGATGTTCTCTTCTTCATG 2820				
QY	63 -----				
DB	2821 CTATAAACTTACATTTTATATAATGTTGTTGTTAGTCAAGATCTACAGGATGTA 2880				
QY	69 ValPheAsnTyrLysAspCysAsnAsnThrIleGluLysThrGlu-----				
	83				
DB	2881 GTTTCAACTATAGAGATTGTATATACAACTTACAGAAACGAGGTTAGTCTTTTCT 2940				
QY	83 -----				
DB	2941 GTTCGTCGACAAATTCGATGCTCATATGTTTCTATGATGATTGTTATTTACTA 3000				
QY	84 -----				
DB	3001 TTTTTCCTGATTGTATGTCAGATGATAGGAGGATTTTCTTCCATTTGCTATGCT 3060				
QY	96 uCysGlySerPheLys-----				
DB	3061 ATGTGTAGCTTCAAGGTGGGCACTATTACAACTGAGGTTTCTCCGGGCGCTTCATA 3120				
QY	101 -----				
DB	3121 TCTAACACTGGAATGCTACTGCTTTTCACTGCTATCTTCACTGTTGTTACT 3180				
QY	101 -----				
DB	3181 ATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3240				
QY	102 -----				
DB	3241 GAGAAACATGTCACAGATGAGCTTACAAACAAATGCTGTGTCTATGACAGGAGCTGC 3300				
QY	104 InPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu-----				
DB	3301 AATTTCAATTTGAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGATGATGTTTGA 3360				
QY	120 -----				
DB	3361 TCGAATTTCTGATTTGCTATGCTTTATTTAATGAGGTTATAGTTAAAAAGGTCCTTTC 3420				
QY	121 -----				
DB	3421 Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136				
QY	137 IlePheGlu-----				
DB	3481 AATTTGAGGCGAGTACTTAACTTAACGTTAATGGAATCCGATAGCTGTTGAAAT 3540				
QY	140 -----				
DB	3541 TTTGTTATATTCATCTTATTTGTTACTAGAAAGAAAGAGATGAGATAAATTTGA 3600				
QY	149 uProPheSerLeu-----				
DB	3601 GCCCTTCTCTCTGTGTGTAACCTCGAAGACCCCTTCGATTAAATACCTTAATAGCAGTAC 3660				
QY	154 -----				
DB	3661 TCCTTGCTTCTTGTGTGTAACCTCTGTAATCCACCAATATGTTTTCAGAGCTCGAA 3720				
QY	156 sProArgLysArgArgGlnArgGlyIleArgAsnAsnThrArgArgLeuLysValCysPhe 176				
DB	3721 ACCTGTGAAGGCGAGCAAAAGAGGTGGCAAAATAACACCGAGAGCTTAAAGTATGCTT 3780				
QY	176 eLeuProLeuAspSerProSerLeuThrAsnGlyIleThrGluAsnGlyIleThrLeuLeuAs 196				
DB	3781 TTTACCGTTGATTCACCCAGTTTAACCTAATGACAGAAATGGAATCACCCCTACTTA 3840				
QY	196 nAsp-----				
DB	3841 TGATGTAAATATCATATCTTCTGTGCGCTTCCTTGAGCTTAGAACTTATATACAG 3900				
QY	197 -----				
DB	3901 AAGAAGATACAATGCGCTGATGTTTGTAGTTTGTACTTCTCTGCAATCTTCTTCCGA 3960				
QY	198 -----				
DB	3961 GGGTATTTGTTACCAAGACTGATGACAAATTAATGAGCATGCTACAGGAAACCGTGGTTT 4020				

QY 202 uclytyrproglualathrgluleuallaglylnphgumethrserasnlepropr 222
 |||||||
 Db 4021 AGGATATCCCGAGCAGACAGACGACTGCTGACAAATTTGGATGACCAGCAATTCACC 4080
 QY 222 oalailaialhisserSerleuaspalaglylalyvalilleuethrSerglualaya 242
 |||||||
 Db 4081 ACCCATAGCCCACTCTCTCTGAGCGCTGCTAAAGTTATATTGACAAAGCAAGCTGT 4140
 QY 242 lvalproalathrlysthrArglyleuSerlialguatgSerglualarg----- 259
 |||||||
 Db 4141 GGTCCCTGCTACTAAGACAAAGAAAGTTATCTGCTGAGCATGACAGGGCTAGAAGGTTGT 4200
 QY 259 ----- 259
 Db 4201 TCATCATGACACCCGTCATCATTAATTAACATTCCGTGTTCACAATGTTCTCTATT 4260
 QY 260 -----Serhisleu 263
 |||||||
 Db 4261 ATGATAGAGTGTATTAGTACTGTCATTAACGAGAAATTTCTCCAGCCACTACT 4320
 QY 263 ulenulnlysarglinphetyrhissertisargyal----- 275
 |||||||
 Db 4321 TCTTCAGAAACGCCCAATTTCTATCATCTTCACAGAGTCAGGTGATCCAAAGTTCCTCACC 4380
 QY 275 ----- 275
 Db 4381 TACTCTTAGCATTTCTTTAAATTCATCATGATGATATCTTATCAAGCATACTGTGT 4440
 QY 275 ----- 275
 Db 4441 TTGTTCTACCAAAATTTGTATTGTGATCTGATGATCAACGCAAAATAGTTATGTCGA 4500
 QY 276 -----Gln-Prometa 279
 |||||||
 Db 4501 TGTGTCTCCGTTTATTGCCACTAACCAAAAAATGATGTTCTGTGCAAGCCATG 4560
 QY 279 lalengluinvalmetseraspargspserglunspglualaspapapvalalala 299
 |||||||
 Db 4561 GCCTGAGCAAGTAATGTGTGACCGGAGTAGCGAGATGAAGTCATGATGATGTTGCGAG 4620
 QY 299 spphgelnasparg----- 303
 |||||||
 Db 4621 ATTGTAAGATCGCAGATTCATCATGATTTCTTCTGCGTTCATTAAATAGACAACA 4680
 QY 304 -----Glnmetleu 306
 |||||||
 Db 4681 AAATGGTATGATGATTAACCTGCTAATGCGCTTTGAAACTTAAAAAGCTGAGATGCTT 4740
 QY 307 ASPASPphevalaspvalasnlyspglulysglinphemethisrleutripsnsrpe 326
 |||||||
 Db 4741 GATGACTTGTGATGTGAATTAAGATGAAAAAGCAATTCATGATCTTTGGAACTGTTT 4800
 QY 327 valarglysglnargyal----- 332
 |||||||
 Db 4801 GTAAGAAAAAAGTA-ACACTTCTCTTACACATGAACAGACAAAAAGACCTTATG 4859
 QY 332 ----- 332
 Db 4860 TCTTACATTCATACCTGCTTAATGATTTGCTTATGGAAGCTTGAGCTCAATATGAT 4919
 QY 333 -----llealaspgllyhisilesertipalacysglualapheser 346
 |||||||
 Db 4920 TGTGATGTTTCAGGGTTATACAGATGCTCATATTTCTTGCGCATGTGAAGCATTTTCA 4979
 QY 347 ArgphetyrglulysgluleunhisarglyrSerSerleuphetip----- 361
 |||||||
 Db 4980 AGATTTCAGAAAAAGATTGCAACCGTACTCATCTCTCTCGTAATATATAGTACACC 5039
 QY 361 ----- 361
 Db 5040 AAACATATACAGACACATACACTACATCAATCTGTTGTTCTGTAATAAATA 5099

QY 362 -----CysTrpargluepheleuilelysleuTrpasnHisglyleuValasp 377
 |||||||
 Db 5100 AAAATTCAGAGTGTGGAGATTTGTTGATTAACATGGAACCATGACTTGTGCAC 5159
 QY 378 SeralathrIleasnascysasnThrilleuengluanCysargasnSerSerapthr 397
 |||||||
 Db 5160 TCAGCCACCATCAACAACGCAATCAATCCCTCGAAGATTGCCGTAATGCTCAGACACC 5219
 QY 398 ThrThrThrAsnAsnAsnSerValaspargProserAspSerAsnThrAsnAsn 417
 |||||||
 Db 5220 ACCACCACCAACAACAACAGTGTGGTGTCCAGTCAAAACCAACAACAT 5279
 QY 418 AsnIleValaspHisProkaspIleasnAsnLysAsnAsnValaspAsnLysAsp 437
 |||||||
 Db 5280 AACATTGTGRTCAATCCCATGACATTAACAACAAGAACAAATGTCACACAGACAT 5339
 QY 438 AsnSerArgaspLysValillelys 445
 |||||||
 Db 5340 AACAGCAGACAAAGTATTA 5363
 RESULT 12
 ATRP5LE2/ 91660 bp DNA linear PLN 22-DEC-2000
 LOCUS Arabidopsis thaliana RPP5 disease resistance locus, main contig.
 DEFINITION AF180942 U97106
 ACCESSION AF180942.1 GI:6449044
 VERSION
 KEYWORDS
 SEGMENT
 SOURCE
 ORGANISM
 Arabidopsis thaliana.
 Arabidopsis thaliana.
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 76020 to 83095)
 2 of 2
 Parker,J.E., Coleman,M.J., Szabo,V., Frost,L.N., Schmidt,R., van
 der Biezen,E.A., Moores,T., Dean,C., Daniels,M.J. and Jones,J.D.
 The Arabidopsis downy mildew resistance gene RPP5 shares similarity
 to the toll and interleukin-1 receptors with N and L6
 Plant Cell 9 (6), 879-894 (1997)
 9212464
 MEDLINE 97355983
 PUBMED 9212464
 REFERENCE 2 (bases 1 to 91660)
 Noel,L., Moores,T.L., van der Biezen,E.A., Parniske,M.,
 Daniels,M.J., Parker,J.E. and Jones,J.D.
 Pronounced intraspecific haplotype divergence at the RPP5 complex
 disease resistance locus of Arabidopsis
 Plant Cell 11 (11), 2099-2112 (1999)
 JOURNAL 20025650
 MEDLINE 10559437
 PUBMED 10559437
 REFERENCE 3 (bases 76020 to 83095)
 Coleman,M.J., Parker,J.E., Szabo,V., van der Biezen,E.A.,
 Daniels,M.J. and Jones,J.D.G.
 Direct Submission
 Submitted (10-Apr-1997) Sainsbury Laboratory, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, UK
 4 (bases 1 to 91660)
 Noel,L., Moores,T., van der Biezen,E.A., Parniske,M., Daniels,M.J.,
 Parker,J.E. and Jones,J.D.G.
 Direct Submission
 Submitted (26-AUG-1999) Sainsbury Laboratory, John Innes Centre,
 Colney Lane, Norwich, Norfolk NR4 7UH, United Kingdom
 On Nov 18, 1999 this sequence version replaced gi:2109274.
 COMMENT Location/Qualifiers
 FEATURES
 source
 1..91660
 /organism="Arabidopsis thaliana"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="4"
 /map="C18a"
 5354..9673
 /gene="La-1"
 /pseudo
 join(5354..5829,5977..7057,7241..7510,7610..8071,

CDS

```

8281..9366,9458..9673)
/gene="La-I"
/note="RPP5 homolog"
/pseudo
repeat_region
15312..20023
/rpt_family="retroelement castor"
/rpt_type="dispersed"
LRR
15312..15431
/gene="encodes gag/pol"
/pseudo
LRR
15904..20023
20048..22736
/gene="La-H"
/pseudo
CDS
15904..20023
20048..22736
/evidence="not experimental"
join(20048..20440,20985..22416,22533..22736)
/gene="La-H"
/note="RPP5 homolog"
/pseudo
gene
25823..35584
/codon_start=1
/note="La-G"
/note="interrupted by retroelement pollux2"
/pseudo
CDS
35225..35899,34108..35235,35366..35584)
/evidence="not experimental"
join(25883,26301,26428..27343,32765..32926,33111..33407,
35225..35899,34108..35235,35366..35584)
/gene="La-G"
/note="RPP5 homolog"
/pseudo
repeat_region
27344..32762
/codon_start=1
/rpt_family="retroelement pollux1"
/rpt_type="dispersed"
LRR
27344..32791
27937..32166
/gene="encodes gag/pol"
/pseudo
LRR
32315..32762
36424..39175
/gene="La-F"
/pseudo
CDS
32315..32762
36424..39175
/evidence="not experimental"
join(36424..37022,37167..38250,38389..38695,38855..39175)
/gene="La-F"
/note="RPP5 homolog"
/pseudo
gene
39222..43288
/codon_start=1
/gene="La-E"
/pseudo
CDS
39222..43288
46918..56304
/evidence="not experimental"
join(39222..39754,39899..40064,42749..43288)
/gene="La-E"
/note="RPP5 homolog"
/pseudo
gene
46918..56304
/codon_start=1
/note="interrupted by retroelement pollux2"
/pseudo
CDS
46918..56304
47393,47538..48635,48820..48990,54412..54573,
54738..55054,55172..55546,55648..55950,56080..56304)
/gene="La-D"
/note="RPP5 homolog"
/pseudo
repeat_region
48991..54409
/rpt_family="retroelement pollux2"
/rpt_type="dispersed"

```

```

LRR
48991..49438
49584..53813
/gene="encodes gag/pol"
LRR
53962..54409
complement(58189..58275)
/misc_feature
/note="similar to 5S ribosomal RNA gene repeat unit"
/note="similar to 5S ribosomal RNA gene repeat unit"
59720..65992
/gene="La-C"
/pseudo
CDS
61136..72604
/evidence="not experimental"
join(68136..68603,68739..69793,69993..70280,70398..70870,
71092..72240,72368..72604)
/gene="La-B"
/note="RPP5 homolog"
/pseudo
gene
68136..72604
/codon_start=1
/note="La-B"
/pseudo
CDS
68136..72604
/evidence="not experimental"
join(68136..68603,68739..69793,69993..70280,70398..70870,
71092..72240,72368..72604)
/gene="La-B"
/note="RPP5 homolog"
/pseudo
mRNA
<77312..>82401
/gene="RPP5"
/evidence="experimental"
join(<77312..77790,77917..78991,79181..79474,79581..79994,
80101..80598,80811..81842,81970..82180,82319..82401)
/gene="RPP5"
/product="RPP5"
join(77312..77790,77917..78991,79181..79474,79581..79994,
80101..80598,80811..81842,81970..82180,82319..82401)
/gene="RPP5"
/codon_start=1
/product="downy mildew resistance protein RPP5"
/protein_id="AA08790.1"
/db_xref="GI:6449046"
/translators="MASSSSGRRYDVEPSPGYDVRKTFSLHLKALDGSINTFI
DHGERSRTIAPELISAIRREARSIYIFSKASVSSWCLNEVEIKCFENDLGMVLP
VYFVDPSEVRKOTGEGKVFETKCEKDGKDPGOKGOKVOALDIDNANGDGLNG
PNEAHWEKISNDVSNKLTIRSKCPDPPGIGIAHIEAKSVICSEKRAHWGIMGOS
GIGKSTIGRALFQSLSOPFHRAFLYKSTSGSDVSGMKLSMKRELLSEITIGODIRI
EHFGVBOIRLNHKKVILLDDVDNLEETKSTSGSDVSGMKLSMKRELLSEITIGODIRI
IDLVEYKLPSCGLAKMISOVARGKSDPPDPKRLAEPVAETLSGLISVYSSTL
KGRKDKWVWMPRIKNDSDKTEETLRYGVRLNKKRDELKPCACPNKSPKVSNTK
ELLEDVGLTMLAESELIIRTPYGYTEMNILEKIGREIDRAKSKCNGKCOPTNPE
DIRVLETKETGETELICIRLPHRGVYTTSPILIDKSEKGMNLOYLTCYMSQVLD
OSIYFPRKRLKPLMNDGPKRLPSNFKAYEYELBPNNSKIEKMDTOTDLSLAKKA
DIYNSYKTRIPDLSTAIINEELNECESLETPDSSIONAKLLELMKMGALLIDELK
SLEKCEKLELSPWSASRECTGATYFFRKIAVWNCPLKRLPSNFKAYEYELK
MEISELEKMDGOSUGSLKEMNLKISNKLKELPDLSIAINLEEDLGCGLVTLDS
SIQNAIKLITLIDMSKCNLESFTYVNLNLSLELIDTGGPNLRNPPATKMGATRLS
KIRLPEGRNEIYEDCFMKNKLPAGLDLIDCLMKCPKCFERSEDLPLNYSGLKLR
LMEGIOSLSLEEMDSESENLKELPDLSKRNKLUCLCSGKSLVLPISNGLVNLN
RLYNKRCOTLEVLPTDVNLSSLETIDLSGSSLNTPFPLISNVCIALENTALELP
DLSKATKLESLLNNCKSLVLPSTIGNQNRRLRYMNRCTGLELLPTDVNLSSLELP
DLSCSSLRFLPLISRIECLVAENVAIEVPCJEDETRLVLMCCOAKNMSIP
IFRLTSLTADFDGCGVAKALSADAVNTMEDHSCVCLSENIETCEBHWACSDY
YSDFEYVNRNPIRLSPMTYNDVDFKFCSTITIKCGVRLVLYOETEBNODTTSKRR
MRMTSGSEEDIMLPGQIIVADTGLAALMTETLSLOGCASSTSLGEGALCVDYDMIN
EEODEQIPILYPIYDIDDMMWSLFGDIDDMMWSLVSAAE"
BASE COUNT 28475 a 17266 c 17533 g 28386 t
ORIGIN
Alignment Scores:
Pred. No.: 8.02e-154 Length: 91660
Score: 1798.50 Matches: 438
Percent Similarity: 43.38% Conservative: 1

```

Best Local Similarity: 43.28% Mismatches: 6
 Query Match: 75.63% Indels: 570
 DB: 8 Gaps: 13
 US-09-890-220-2 (1-445) x ATPP5LE2 (1-91660)

QY 1 MetcysArgInAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 88263 ATGTCTAGGCAAAATGTCGGCCGAATCCTCAGCGGAGGATGATTCACTCATGAG 88204
 QY 21 AsnLeuLeuIleTyrCysLysProValAlaArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 88203 AATCTTGATGATGATTGTAACCTGCTGACATATATACATCTTCACCTCGCTCTCA 88144
 QY 41 GlyAsp----- 42
 DB 88143 GGCAACGATGATTTGGCTTCTCTCTCATCATATATAGCTCAGTAATCTTCATCTCTG 88084
 QY 42 ----- 42
 DB 88083 TGTAGATCACCCACTAATAGTTGAGTTTGTGTAAGCTGATTTGGTGTGATTCATGGCGA 88024
 QY 43 -----ProSerPhe 45
 DB 88023 GTGTGTGCTTCTTTTGTCTCTCAAAATTTGAACCTGTGTGTGTGTGTCAGCCATCGTTT 87964
 QY 46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysSer----- 62
 DB 87963 CTTCAGATGCTTGAACCTACCAAAATGGAGCAAGCGCAAAAGAAAGTA-TGTTTCTT 87905
 QY 62 ----- 62
 DB 87904 CTGGAATGAGCTGTACAGTATATGTTATTATCTTACTCTAATATGAGAGCGTAG 87845
 QY 62 ----- 62
 DB 87844 ACCATTTTATCTTTGTTGAGTAGATATGACATATGATGTTTCTTGTGTCATG 87785
 QY 63 -----ArgSerThrGlyMetVal 68
 DB 87784 CTATTAACCTTACATTTTATAAATGTTGTTGGTTAGTCAAGATCTGAGATGGA 87725
 QY 69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 DB 87724 GTTTCACTAATAAGATTGTATATAACATTTACAGAAACGTAAGGTTAGCTTTTCT 87665
 QY 83 ----- 83
 DB 87664 GTTCGTGCAAAATTCGATGTCATGCTATGTTTCTTCTAGATATTTGTTATTACTA 87605
 QY 84 -----ValArgGluAspCysSerCysProPheCysSerMetIle 96
 DB 87604 TTTTTCCTGATTTGTCTATGACGTTAGGAGGATGTTCTTCCATTTTGTCTATGCT 87545
 QY 96 uCysGlySerPheLys----- 101
 DB 87544 ATGTGTAGCTTCAAGGTGGGCAACTATTACAACGAGGTTCTTCCGGGGGCTTTCATA 87485
 QY 101 ----- 101
 DB 87484 TCTAACACTGTGAAATGCTACTGCTGTTTCATGCTATATCTTCACTGTTGGTTACAT 87425
 QY 101 ----- 101
 DB 87424 ATTTTGTGTTTGTGTTGTTGTTCTTCTCACTTTTTCGAACCTGAGTGTGCTTATCT 87365
 QY 102 -----GlyLeuG 104
 DB 87364 GAGAAACATGTCAGATGAGGCTTACAACCAATTTCTGTGCTATAGCAGGGGCTGTC 87305
 QY 104 InPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu----- 120
 DB 87304 AATTCATTTGAAATTCATCTCATGATTTATTTGAATTTGCAAGGATATGTGTTTAA 87245

QY 120 ----- 120
 DB 87244 TGGAAATCTTGATTTGCTATGCTTATTAATGAGTTATGTTAAAAAGGCTCTTC 87185
 QY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136
 DB 87184 CTATGTAGCTTTCGGAAGAAATACAGACAGTAAATGTTTCTGTAACCTTAATTCCTTC 87125
 QY 137 IlePheGluGlu----- 140
 DB 87124 AATATTGAGT-CAGTAACCTTAACTTGTTAAATGGGAATCCGATAGCTGTGAAA 87066
 QY 141 -----GluGlySerAspAspLysPhe 148
 DB 87065 TTTTGTATATTCATCTTATTTGTACTAGAAAAGAGATGATGCAATAAATTT 87006
 QY 149 GluProPheSerLeu----- 153
 DB 87005 GAGCCCTCTCTCTGTAACCCCTCAGAAACCTTCGATTAATACCTTAATAGCAGTA 86946
 QY 154 -----CysSer 155
 DB 86945 ACTCCTGCTCTCTGTCAGTACATCTCTGTAATCCAAACCATATGTTTTCAGCTCG 86886
 QY 156 LysProArgLysArgArgGlnArgGlyArgAsnThrArgArgLeuLysValCys 175
 DB 86885 AATCTCTGTAACCGGAGACAAAGAGTGGCAGAAATACACAGGAGCTTAAAGTATGC 86826
 QY 176 PheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyLleThrLeuLeu 195
 DB 86825 TTTTACGCTTGATTTACACCCAGTTTAACATATGACAGAAATGGAATCACCCCTACT 86766
 QY 196 AsnAsp----- 197
 DB 86765 AATGATGTGTAATATCATATCTTCTGTGCGTCTCTGCTTGAACCTTCAATATAC 86706
 QY 197 ----- 197
 DB 86705 ACAAGACATCAATAGCCCTGATTTGTTGTTTGTACTTCTCTCGCATCTCTTTCG 86646
 QY 198 -----GlyAsnArgGly 201
 DB 86645 GAGGATGTGTACACAGAACGTGATGACAAATTAATGCGATCTACAGAAACCGTGT 86586
 QY 202 LeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIlePro 221
 DB 86585 TTAGATATCCCGAGGCAACAGAACCTGTGACAAATTTGAGATGCCGCAACATTTCCA 86526
 QY 222 ProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAla 241
 DB 86525 CCAGCCATAGCCCACTCTTCTGTGACGTGTGTCTTAAGTATATGCAAGCGAAGCT 86466
 QY 242 ValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
 DB 86465 GTGTGCCCTGTACTAAGACAGAAAGTATCTGTGACGATCAGAGGCTGAAGGTTT 86406
 QY 259 ----- 259
 DB 86405 GTTCATCATGACCCCGTCATCATTAATTAACATTCCTGTGTTTACAAATGTTCTCTTA 86346
 QY 260 -----SerHisLeu 262
 DB 86345 TTATGATATAGCTTTATATGATCTGCAATTAATTAACGAGAAATTTCTTCAGCCACCTA 86286
 QY 263 LeuLeuGlnLysArgGlnPheTyrHisSerHisArgVal----- 275
 DB 86285 CTTCCTGAAAGCGCAATTTATCATCTCTACAGAGTCCAGGTGATCCAACTTCCTTCA 86226
 QY 275 ----- 275
 DB 86225 CTTACTTCTAGGATTTTCTTTAAATGCTCATGATGATATCTTATCAAAAGCATACTTG 86166

OY 275 ----- 275
 Db 86165 GTTGTTCATCCAAATTTGATTTGATGATGATCAACGCAAAATAGTTATGTC 86106
 OY 276 ----- Gln-Prom 278
 Db 86105 CATGTCTCTCCGTTTATGTCACCAAAATGACGTTCTGTGCAAGCCAT 86046
 OY 278 TAlAlEuGlInValMetSerAspArgSerGluAspGluValAspAspValAl 298
 Db 86045 GGGCTTGACCAAGTATGTCGACCGGATGACGATGATGATGATGATGATGTC 85986
 OY 298 AAspPheGluAspArg ----- 303
 Db 85985 AGATTTTGAAGATGCCAGGATATTCATGATTTCTGCGTCAATTAATAGACA 85926
 OY 304 ----- GlnMet 306
 Db 85925 GAAATGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 85866
 OY 306 EAspAspPheValAlAspValAlAspGluValPheMetHisLeuThrAsnSerP 326
 Db 85865 TTGATGATCTTGATGATGATGATGATGATGATGATGATGATGATGATGAT 85806
 OY 326 heValAlArgLysGlnArgVal ----- 332
 Db 85805 TTGTAAAGAAAACAAAGTA-ACACTTCTTACACATGACACACAAAGAACCTTA 85747
 OY 332 ----- 332
 Db 85746 TGCTTACATTCATCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 85687
 OY 333 ----- IleAlAspGlyHisIleSerThrAlaCysGluAlaPheS 346
 Db 85686 ATGTGTGATGTTTCAGGCTTATGATGATGATGATGATGATGATGATGATGAT 85627
 OY 346 eArgPheArgLysGluValLeuHisArgTyrSerSerLeuPheThr ----- 361
 Db 85626 CAAGATTTTACGAGAAAGAGTTGACACCGTACTCATCTCTCTCTCTCTCTCT 85567
 OY 361 ----- 361
 Db 85566 CCAACATATACAGACATTAACATCACTATCACTCTCTCTCTCTCTCTCTCT 85507
 OY 362 ----- CysTrpArgLeuPheLeuIleLysLeuThrPheHisGlyLeuVal 377
 Db 85506 TAAAAATTTCCAGGTGTTGATGATGATGATGATGATGATGATGATGATGAT 85447
 OY 377 sPseAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 397
 Db 85446 ACTCAGCCACATCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 85387
 OY 397 hTrThrPheAsnAsnAsnAsnSerValAspArgProSerAspSerSerPheAsn 417
 Db 85386 CCACACACACACACACACACAGTGTGATGTCACAGTACACAAACACAAACA 85327
 OY 417 sAsnIleValAlAspHisProAsnAspIleAsnAsnLysAsnValAlAspAsn 437
 Db 85326 ATAACATTTGAGTCTCTCCATGATGATGATGATGATGATGATGATGATGAT 85267
 OY 437 sAsnSerArgAspLysValIleLys 445
 Db 85266 ATAACAGCAGACAGCAAGTAATTA 85241
 RESULT 13
 LOCUS AX032893 5895 bp DNA linear PAT 21-SEP-2000
 DEFINITION Sequence 6 from Patent WO0044918.
 ACCESSION AX032893
 VERSION AX032893.1 GI:10279828
 KEYWORDS
 ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 5895)
 AUTHORS Dean, C. and Gendall, A.
 TITLE Methods and means for modification of plant characteristics using
 the vernalization gene vrn2
 JOURNAL Patent: WO 0044918-A 6 03-AUG-2000;
 DEAN CAROLINE (GB) ; GENDALL ANTHONY (GB) ; PLANT BIOSCIENCE LTD
 (GB)
 FEATURES
 source Location/Qualifiers
 1..5895
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 BASE COUNT 1742 a 1027 c 1070 g 2056 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,44e-142 Length: 5895
 Score: 1659.00 Matches: 424
 Percent Similarity: 42.25% Conservative: 12
 Best Local Similarity: 41.86% Mismatches: 577
 Query Match: 69.76% Indels: 14
 DB: 6 Gaps: 14
 US-09-890-220-2 (1-445) x AX032893 (1-5895)
 OY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 Db 1905 ATGTGTAGGCAAGATTTGTCGCGCAAAATCTCCACCGAGAGAGTATTCACATGAG 1964
 OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuThrAsnIlePheHisLeuArgSerLeu 40
 Db 1965 AATCTCTGATATATGTGAACCTGTCGACTATATACATCTTCCACCTGCTCTCTA 2024
 OY 41 GlyAsn ----- 42
 Db 2025 GCCAA-CGTATGATTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2083
 OY 42 ----- 42
 Db 2084 TGTGTAGATCACCCACTAATAGTTGAGTTGCTAAGCTGATGATGCTGATGATGATG 2143
 OY 43 ----- Pro 43
 Db 2144 GAGTGTGCTTCTTTGTCCTCAATGTAATTTGAACCTGTTGTTGTCGACGCA 2203
 OY 44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysSerSer 62
 Db 2204 TCGTTTCTGCCAAGATGCTTGAACCTACCAAAATTTGGGGCAAAAGGCAAAAGTA-TGC 2262
 OY 62 ----- 62
 Db 2263 GTTTCCTTCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2322
 OY 62 ----- 62
 Db 2323 GCTGTAGACTATTATCTTTGTTGAGTACATATGACATATGAATGATGATGATGAT 2382
 OY 63 ----- ArgSerThrGly 66
 Db 2383 TTCATGCTATCACTTATATTTTACAAAATTTGTTTCTTCTAGTCAAGATCACTGGG 2442
 OY 67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGluLysThrGlu ----- 83
 Db 2443 ATGTGATTTTCAACTATTAAGATTTGTAATTAATACATTAACAAAGAGAGAGTATGCT 2502
 OY 83 ----- 83
 Db 2503 TTTCTGTCTTCGACAAAATTCGATGATGATGATGATGATGATGATGATGATGAT 2562
 OY 84 ----- ValArgGluAspCysSerCysProPheCysSe 94

Db 2563 TTAATATTTTCTGTATGTCACGAGTTAGGAGGATGTTCTGTCATTTTGCTC 2622
 QY 94 rMetLeucGylSerPheLys----- 101
 Db 2623 TATGCTATGTGTACTTCAAGGTGGCAACTATTACACTGAGTTTCTCCGGGCT 2682
 QY 101 ----- 101
 Db 2683 TTCAATCTAACACTGTGAATGCTACGCGTTTAATGTAATACTTTCACGTGTGG 2742
 QY 101 ----- 101
 Db 2743 TTACATATTTTGTGTTGTTGTTGTTCTCTGCTTTTAACTGCTGAGTGTGC 2802
 QY 102 -----G 102
 Db 2803 TTATCTGAGAAACATGTTCCAGTTCCAGTTTACATCCATGCTGTCTATGCAAG 2862
 QY 102 LLeuGlnPheH1sLeuAnsSerHisAspLeuPheGluPheGluPheLysLeuPhe- 121
 Db 2863 GCGTCGCAATTCATTTGATTCATCTCATGATTTATTGAATTGAGTTCAAGTATG-T 2921
 QY 121 ----- 121
 Db 2922 GGTTTTANGAATTTCTGTTTGCCATGCCGTAGTAAAGAGGTTATAGTTAAAG 2981
 QY 122 -----GluGluTyrGlnThrValAsnValSerValLysLeu 134
 Db 2982 GGTCTTCCCTATGTAGCTTTTGGAGAAATACAGACGTTATGTCTGTAAACCTTA 3041
 QY 134 snSerPheL1ePheGlu----- 139
 Db 3042 ATTCCTCATATTGAGTCAAGTTACTTAACTGGTAAATTGGGAATCCTATAGCTG 3101
 QY 140 -----GluGluGlySerAspAsp 146
 Db 3102 GTGAAATTTGTTTATATTCATCCCTATTGTACTAGGAAGAAAGATGATGAT 3161
 QY 147 LysPheGluProPheSerLeuCys----- 154
 Db 3162 AAAATTTGAGCCCTTCTCTCTCG-GTAACTCTCAGAACCCCTTGATTAAATCCTTAATA 3220
 QY 154 ----- 154
 Db 3221 GCAGTAACCTCTGCTTTTCTTGTCAGTACTTCTATATAATCCAAACACATGTTTGC 3280
 QY 155 ---SerLysProAlaGlyArgArgGlnArgGlnArgLysArgLysLeuLys 173
 Db 3281 AGCTCGAAACCTCGTAACCGTAGACMAAGAGGTGGCAAAATAACACAGGAGACTTAA 3340
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyThr 193
 Db 3341 GTATGCTTTTACCTTGATTCACCCAGTTTACTAATGGCAGAAAATGGAATTGCC 3400
 QY 194 LeuLeuAsnAsp----- 197
 Db 3401 CTGCGATGATGATGTAATAATACATCTTCTGTGTATGCTGTGGCTTAGAACATTC 3460
 QY 197 ----- 197
 Db 3461 ATTTTACAGAAAGATACAAATGTCCTGATTGTTTGTATCTCTCTGCAATTC 3520
 QY 198 -----GlyAs 199
 Db 3521 TTCTTGAGGAGTATGTTTACAGAACTGATGTACAAATTAATGCAATGCTACAGGAAA 3580
 QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetLysSerAs 219
 Db 3581 CCGTGCTTAGGATATCCGAGGCAACAGACTTCTGAGCAATTTGAGATGACTAGCAA 3640
 QY 219 nIleProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 3641 CATTCACCGACCATAGCCCACTCTTCTGTGACGCTGTGCTAAAGTTATATTAACAAC 3700

QY 239 rGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaAr 259
 Db 3701 CGAAGCTGTGCTCCCTGCTACTAGACAGAAAGATTATCTGTGAGCCATCAGAGGCTAG 3760
 QY 259 9----- 259
 Db 3761 AAGTTGTTCATCATGACACCCCGTCATCATTAATACCAATCTGTTACAAATGTT 3820
 QY 260 -----Se 260
 Db 3821 CTTCCTAATAGATAGTGTGTTACTGTACTGCCATTAATTAACGAGAAATTTCTTCAG 3880
 QY 260 rHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 3881 CCACCTACTTCTTCGAAACGCCAATTTCTATCATCTCTACAGAGTCCAGTGATCCAGT 3940
 QY 276 ----- 276
 Db 3941 TCCTTCACCTACTCTCTAGGCATTTTCTTAAATTGCTCAGATGATATCTTATCAAGC 4000
 QY 276 ----- 276
 Db 4001 ATACTGTGTTGTTCTCATCTAAATTGTGATTTGATTCTGTATGATCAACGCAAAAA 4060
 QY 276 ----- 276
 Db 4061 ATTATGTCCATCTGTCTCCGTTTATTATGCCACTAACCAAAACATGATTTCTGTGCA 4120
 QY 277 ---ProMetAlaLeuGlnGluAlaMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 4121 CAAGCCAATGCGCGCTTGAGCAAGTAATGTCTGATGCGATAGCGAGATGAAGTCGATGA 4180
 QY 295 pAspValAlaAspPheGluAspArg----- 303
 Db 4181 CGATGTTCAGATTTTGAAGATGCGCAGGTATTCATGATTTCTTGTGCTTCATTAAG 4240
 QY 303 ----- 303
 Db 4241 TAGGCAACAGAAATGATATAGATGTAATCTGTAATGCGTTTGAACCTTAAAAAGC 4300
 QY 304 --GlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuT 323
 Db 4301 TGCAGATGCTGATGACTTTTGATGGAATTAAGATGAAGAAACCAATTCATGATCTTT 4360
 QY 323 rPAsnSerPheValArgLysGln----- 330
 Db 4361 GGAAGCTGTTGTGAAGAAACAAAGTAACACTTCTCTTACACTTGAAACACACAAAA 4420
 QY 330 ----- 330
 Db 4421 AGACCTTATGCTTACATTCATACCTGTCTAAATGATTTGCTTATGGAACCTTGAGCT 4480
 QY 331 -----ArgValIleAlaAspGlyHisL1eSerTrpAlaCysGlu 343
 Db 4481 CAATTAATGATGATGTTTGAGGCGTGTAGAGATGTGCATATCTGTGGCAATGTA 4540
 QY 344 AlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 4541 GTATTTCAAGATTTTACAGAGAAAGTTCAGTCTTACTCATCACTCTCTGGAATAT 4600
 QY 361 ----- 361
 Db 4601 AAGTACACAAACATATACAGACATACATACATCAATTTTGTTCGTTTTCGA 4660
 QY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisG1 374
 Db 4661 AAGAAATAATAAAATTCAGGTGTGAGATGTTTGTGATTAACATATGAAACCAAG 4720
 QY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
 Db 4721 ACTGTGCACTCAGCCACCATACACACTGCAATTCATCTCGAATTTGCCGTAATATC 4780

OY 121 ----- 121
 Db 26416 GGTATTATGAAATTTCTGTTTGGCTATGCCGTTAGTAAGAGTTATAGTAAAG 26475
 OY 122 ----- GluGluTyrGluThrValAsnValSerValIysLeu 134
 Db 26476 GGTCTTCCTCATTTGAGCTTTGGAAAGATACAGACAGTTATGTTCTGTAAGACTTA 26535
 OY 134 snSerPheIlePheGlu----- 139
 Db 26536 ATTCTTCATATTGAGGTCAGTACTTAAACTTGTTAATTGGGAAATCCTATAGCTG 26595
 OY 140 ----- GluGluGlySerAspAsp 146
 Db 26596 GTGAAATTTGGTTATATTCATCCTTATTGTACTAGAGAAAGAAAGATGATGAT 26655
 OY 147 LysPheGluProPheSerLeuCys----- 154
 Db 26656 AAATTGAGCCCTTCTCTCTG-GTAACTCTAGAACCCCTTGATTAATACCTTAATA 26714
 OY 154 ----- 154
 Db 26715 GCAGTACTCCTTCTGTTTCTGTCAGTACTTCTGTATTAATCCAAACCAATGTTTGC 26774
 OY 155 ---SerLysProAlaTyrGlnArgAlaGlnArgGlyIleArgAsnAsnThrArgArgLeuLys 173
 Db 26775 AGCTCGAAACCTCTGAAGCGTGACAAAGAGGTGGCAGAAATTAACACAGAGACTTAA 26834
 OY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThr 193
 Db 26835 GATAGCTTTTACCGTTTGATTCACCCACTTTAGCTAATGACACGAAATAGATGCC 26894
 OY 194 LeuLeuAspAsp----- 197
 Db 26895 CTGCTGAATATGTAAATACACATCTTCTGTGATTCGTTGGCTTGAACATTC 26954
 OY 197 ----- 197
 Db 26955 ATTTTACAGAAAGATACAAATGCTGATGTTTGTAGTTTGTCTCTCCGCATTC 27014
 OY 198 -----GlyAs 199
 Db 27015 TTCTGTGAGGTAATGTTACAGAACTGATGTACAAATTAATGCGTCTACAGGAAA 27074
 OY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetTrpSerAs 219
 Db 27075 CCGTGTGTTAGGATATCCGAGGCAACAGACCTGCTGACAAATTTGAGATGACACAA 27134
 OY 219 nileProAlaIleAlaIleSerSerLeuAspAlaGlyAlaLysValIleLeuTrpSe 239
 Db 27135 CATTCACACGACATAGCCACACTCTCTGAGCGCTGGCTGTTAAAGTTATTAACAC 27194
 OY 239 rgAlaIleValIleProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAla 259
 Db 27195 CGAAGGTGTGTCCCGTACTTAAGCAAGAAAGTTATCTGCTGAGCATCAGAGCTAG 27254
 OY 259 g----- 259
 Db 27255 AAGGTTTGTTCATCAGACACCCCGTCATCAATTAATACCAATCTGTTGTACAATGTT 27314
 OY 260 -----Se 260
 Db 27315 CTTCCTATTATGATAGTGTACTGTACTGCCATTAATTAACCGAGAAATTTCTTCAG 27374
 OY 260 rHisLeuLeuLeuGlnLysArgLlnPheTyrHisSerHisArgValGln----- 276
 Db 27375 CCACCTACTTCTTCAGAAAGCCCAATTTATCATCTCACAGAGTCCAGGTGATCAAGT 27434
 OY 276 ----- 276
 Db 27435 TCCTGACCTACTTCTTAGGCAATTTCTTAATTTGTCATGATGATCTTATCAAGC 27494

OY 276 ----- 276
 Db 27495 ATACTGTGTTCTTCATCTAAATTTGTATTTTGATTCGTGATGATTCACGCAAAAA 27554
 OY 276 ----- 276
 Db 27555 ATTAGTCCAGTCTCTCCCTTTTATGGCACTAACCAAAAACGTCATGTTCTTGGA 27614
 OY 277 ---PromElalaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 27615 CAAGCCAAATGGCGCTTGAGCAAGTAATGTCTGATCGGATGAGGAGATGAAGTCATGA 27674
 OY 295 pasPValAlaaspPheGluAspArg----- 303
 Db 27675 CGATGTGCAGATTTTGAAGATCCCGAGATTCATGATTTCTTCTCGTTCATTAAG 27734
 OY 303 ----- 303
 Db 27735 TAGCCACAGAAAAATGATAGATGATGATGATGATGATGATGATGATGATGATGAT 27794
 OY 304 ---GluMetLeuAspAspPheValAspValAsnLysAspGlyLysGlnPheMetHisLeu 323
 Db 27795 TGCAGATGCTGTGATGATCTTGTGATGATGATGATGATGATGATGATGATGATGAT 27854
 OY 323 rPasSerPheValArgLysGln----- 330
 Db 27855 GGAATGCTTTGTAAAGAAACAAAGGTAATGATGATGATGATGATGATGATGATGAT 27914
 OY 330 ----- 330
 Db 27915 AACCTTATGCTTACATTCATCTCTGTAATGATGATGATGATGATGATGATGATGAT 27974
 OY 331 -----ArgValIleIleAspGlyHisIleSerTrpAlaCysGlu 343
 Db 27975 CAAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 28034
 OY 344 AlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 28035 GATTTTCAAAATTTTACAGAAAGAGTCTGATGATGATGATGATGATGATGATGAT 28094
 OY 361 ----- 361
 Db 28095 AAGTACACCAAAATATACAGACATTAATGATGATGATGATGATGATGATGATGAT 28154
 OY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlu 374
 Db 28155 AAGAAATTAATAAAATTCACAGGTGTGATGATGATGATGATGATGATGATGATGAT 28214
 OY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe 394
 Db 28215 ACTGTGACACACGAC 28274
 OY 394 rSerAspThrThrThrAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsnAsn 414
 Db 28275 CTCA-----GTCACTAACACACACACACACACACACACACACACACACACAC 28338
 OY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
 Db 28339 CAACACACATTAACATTTGGATCTCCGAAATGACATTAATAAACAGACAAATGTTGACA 28388
 OY 434 nLysAspAsnAsnSerArgAspLys 442
 Db 28389 CAAGCACAATTAACAGACAGACAG 28413
 OY 28389 CAAGCACAATTAACAGACAGACAG 28413
 Db 28389 CAAGCACAATTAACAGACAGACAG 28413
 RESULT 15
 AFPCAT 201471 bp DNA linear PLN 29-JUN-1999
 LOCUS Arabidopsis thaliana DNA chromosome 4, ESSA I FCA contig fragment
 DEFINITION No. 7.
 ACCESSION 297342
 VERSION 297342.2 GI:5302800
 KEYWORDS Arabidopsis thaliana.

ORGANISM	Arabidopsis thaliana	CDS	/gene="d14415w" join(12153..12312,12416..12714,12983..13062,13151..13466) /gene="d14415w" /note="contains EST gb: 1216777.935272;906153;947895;315626" /codon_start=1 /product="DNA-binding homeotic protein Atb-2" /protein_id="CAB10452.1" /db_xref="GI:2245033" /db_xref="SWISS-PROT:005466" /translation="MWFEDDGLSLGLNFPKKQINLKSNSPSVTPSSSSGGLRRS SWNESTSSVPNSDSQKRTFRIGIDVNRPSIAEYGEDADAGVSSPSVSSGK RSEREDTDPQCSRGISDDEGDSKRLRSKDSALILETFKHSTLNPQKOLA KOLGLARQVEVWFQNRARATKLTQCEVCEFLRCCENTLLENRLQREVELALK LSQRYMHSPPPTTLTMCSCSHVSPPPQQAASAHHRSLPVANMAFATISGLT FPAIRRS"
REFERENCE	1 (bases 1 to 201471) Bevan, M., Stiekema, W., Murphy, G., Wambutt, R., Pohl, T., Terry, N., Kreis, M., Kavanagh, T., Entlan, K.D., Rieger, M., James, R., Puidomenech, P., Hatzopoulos, P., Obermaler, B., Duesterhoff, A., Jones, J., Palme, K., Ansgore, W., Delseny, M., Bancroft, I., Mewes, H.W., Schueller, C. and Chalhatzis, N.	exon	12153..12312 /gene="d14415w" /number=1
JOURNAL	Unpublished	Intron	12313..12415 /gene="d14415w" /number=1
REFERENCE	2 (bases 1 to 201471) EU Arabidopsis sequencing project.	exon	12416..12714 /gene="d14415w" /number=2
AUTHORS	Submitted (25-JUN-1999) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopfersplitz 18a, D-82152 Martinsried, FRG, E-mail: schnelleemips.biochem.mpg.de,mayeremips.biochem.mpg.de Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bsrc.ac.uk	Intron	12715..12982 /gene="d14415w" /number=2
COMMENT	On Jun 30, 1999 this sequence version replaced gi:2245031. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ this fragment has an overlap with AFCA6 at the 5' end and an overlap with AFCA8 at the 3' end.	exon	12983..13062 /gene="d14415w" /number=3
FEATURES	Location/Qualifiers	Intron	13063..13150 /gene="d14415w" /number=3
source	1..201471 /organism="Arabidopsis thaliana" /variety="Columbia" /db_xref="taxon:3702" /chromosome="4" 1..8956 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid EM16B10.IB5K23" 21108..37844 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid EM16B10.IB6C5" 37120..39085 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid EM16B10.IB4A23" 37904..128469 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="BAC IGF5D3" 38368..63818 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid CC3M3" 106243..128469 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid CC16N19" 113141..132730 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid CC20I2" 129847..157017 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid CC1B24" 152677..190349 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid G16599" 185303..201471 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /clone="cosmid G11319" 12153..13466	exon	13151..13466 /gene="d14415w" /number=4 /gene="d14420c" 14609..16030 complement(14609..16030) /gene="d14420c" complement(14609..16030) /note="sequence differences between cDNA and gDNA similarity to hydroxyproline-rich glycoprotein (clone HYP3.6) kidney bean (fragment)" /codon_start=1 /product="glycoprotein homolog" /protein_id="CAB46042.1" /translation="WEARSIKKPIOLGKKNEDONPKKYSRFTFKALITLVCAVVP FISQPELANQTRLELHLVFGTAVSGILSRRYDGGGGGGSNSGHRKADSHN NSHSTVPKILVSVFNGHSESEPSDSSGQRFQTKKNTKHKTIPEVERFVDR VSENNKPKLLPVRSLNRSVSDSGNSGKWKVRSRELKTLGDNSDVLPSPI PKRSRSSSSSSSSKVESLPSVKNTLVESQPLKNTLPSPSSFSRPSGNDIPNLS EPHSPPEPPPPPLPAFYNSSRKDHGIVYVERESSVHKTRKAGSGFPPEPP PPPEVYVKSPPTKRFLSNERRKSSQKKRRAPKVKVMSDPLVESKEQDTKQORS NIGSKAVESSENGQKRGNEIHDEVEKTIVEEGVSEINSGSDVKKADEFLAKFRE QIRLQRIEIKRSTNKISANSSR" complement(14609..16030) /gene="d14420c" /number=1 16148..16170 /note="23bp T tandem repeat" 17691..19115 /note="CDNA1KMY" 17793..19171 /gene="d14425c" complement(join(17793..18046,18147..18250,18348..18399, 18481..18571,18886..18960,19058..19171)) /gene="d14425c" complement(join(17793..18046,18147..18250,18348..18399, 18481..18571,18886..18960,19058..19171))
source	gene	repeat_region	16148..16170 /note="23bp T tandem repeat"
source	gene	misc_feature	17691..19115 /note="CDNA1KMY"
source	gene	gene	17793..19171 /gene="d14425c" complement(join(17793..18046,18147..18250,18348..18399, 18481..18571,18886..18960,19058..19171)) /gene="d14425c" complement(join(17793..18046,18147..18250,18348..18399, 18481..18571,18886..18960,19058..19171))
source	gene	CDS	12153..13466

QY 154 ----- 154
 Db 40886 GCAGTAACCTCGCTTTCTCTGTCAGTACTTCTCTATAAATCCACACCAATGTTTGC 40945
 QY 155 ---SerLysProAlaGlySerArgArgGlnArgGlyGlyArgGlnAsnThrArgArgLeuLys 173
 Db 40946 AGCTCGAAACCTCGTACGACGTAGACAAAGAGGTGGCAGAAATATACACACGAGGACTTAA 41005
 QY 174 ValLysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyLysLeuThr 193
 Db 41006 GTATGCTTTTACCGTGTGGATTCACCCAGTTTACGCTAATGGCAGAAATGGAATTCGCC 41065
 QY 194 LeuLeuAsnAsp----- 197
 Db 41066 CTGGGATGATGGTAAATACACATCTCTCTGTGATTCGTTGCGTTAGAACCTTC 41125
 QY 197 ----- 197
 Db 41126 ATTTTACAGAAAGATACAAATGTCCTGATGTTTGTAGTTTGTACTTCTCTCCGCAATTC 41185
 QY 198 -----GlyAs 199
 Db 41186 TTCTTTCGAGGCTAATGTTTACCAGACTGATGTACAAAATTTAATGGCATGCTACAGGAAA 41245
 QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
 Db 41246 CCGTGTGTAGGATATCCCGAGGACACAGAGCTTCTGACATTTGAGATGACAGCA 41305
 QY 219 nIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 41306 CATTCACACGACGACATAGCCCACTCTCTGTCGACGCTGTGCTAAAGTTATATTAACAC 41365
 QY 239 rGluAlaValAlaProAlaThrAlaThrArgLysLeuSerAlaGlyLysSerGluAlaAr 259
 Db 41366 CGAAGCTGTGGTCCCTGCTACTACAGACAGAAAGTTATCTGTGAGCGATGAGAGCTTAG 41425
 QY 259 g----- 259
 Db 41426 AAGGTTGTTCATCATGACACCCCGCTCATCATTAATTAACCATCTGTTTACAANAATGT 41485
 QY 260 -----Se 260
 Db 41486 CTTCCTAATTAATGATAGTGTACTGACTGCTGCATATTAACCGAGAAAATTTCTTCAG 41545
 QY 260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 41546 CCACCTACTTCTTCGAAACCGCAATTTCTATCATTTCTACAGAGTCCAGGATCCAAAGT 41605
 QY 276 ----- 276
 Db 41606 TCCTTCACCTACTCTCTAGGCAATTTCTTAATTGTCTCATGATGATATCTTATCAAGC 41665
 QY 276 ----- 276
 Db 41666 ATACTGTGTTGTCTCATCTAAATTTGTATTTGATTCGTATGTATCAACGCAAAAAA 41725
 QY 276 ----- 276
 Db 41726 ATTAATGTCATGTTGTCTCCGTTTATTTATGCCACTACCAAAAATGCAATGTTTCTTGCA 41785
 QY 277 ---PrometAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 41786 CAAGCCATGGGGCTTGAGCAAGTAATGTGATCGGATAGCGAGATGAAGTGCATGA 41845
 QY 295 pAspValAlaAspPheGluAspArg----- 303
 Db 41846 CGATGCTTCAGATTTTGAAGTCGCCAGATATTCATGATTTCTTTCTGCGTTCAATTAAG 41905
 QY 303 ----- 303
 Db 41906 TAGGCACAGAAAATGATATACGATGTAATGCTTAATGGCTTTTGAACACTTAAAAAAGC 41965
 QY 304 --GlnMetLeuAspAspPheValAspValAsnLysAspGlnPheMetHisLeuT 323

Db 41966 TGCAGATGCTTGATGACTTTGGGATGGATTAAGATGAAGAAAGCAATTCATGATCTTT 42025
 QY 323 rPasnSerPheValArgLysGln----- 330
 Db 42026 GGAAGCTGTTGTGAAGAAACAAAGTAATCTTCTGTACATTTGAACACACACAAAA 42085
 QY 330 ----- 330
 Db 42086 AGACCTTATGCTTACATTCATTCATACCTGTCTTAATGATTTGCTTATGAAACTTTGACCT 42145
 QY 331 -----ArgValIleAlaAspGlyHisIleSerTrpAlaCysGlu 343
 Db 42146 CAATTAATGATGATGTTTGCAGGGTTATAGAGATGGCATATCTTGGGCAATGTAA 42205
 QY 344 AlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 42206 GTATTTTCAGATTTTACAGAGAAAGTGGCACTGTACTCATCATCTCTCTGGTAATAT 42265
 QY 361 ----- 361
 Db 42266 AAGTACACCAACATATACAGACATTAACACTAATCAATTTTGTTCGTTTCTGCA 42325
 QY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGly 374
 Db 42326 AAGAAATATAAAAATTCAGAGTGTGGAGATGTTTGTATTAACATATGAAACCATGG 42385
 QY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAsnSe 394
 Db 42386 ACTTGTGACATCGACGACCATTCMACACTGCATATCCATCCGAGAAATTTGCCGTAAATAC 42445
 QY 394 rSerAspThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsnTh 414
 Db 42446 CTCA-----GTCATTAACACACACACACAGTGTGGATCTCCAGTACTCAAAAC 42499
 QY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs 434
 Db 42500 CAACAACAATTAACATTTGGATCATCCGATGACATAAAAAACAAGAACATGTTGACAA 42559
 QY 434 nLysAspAsnAsnSerArgAspLys 442
 Db 42560 CAAGGACATTAACAGACAGACAAAG 42584

Search completed: June 20, 2003, 00:55:20
 Job time : 3851 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 22:08:48 ; Search time 330 Seconds

(without alignments)
3036.786 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378
Sequence: 1 MCRONCRAKSSPEEYISDNE.....INKNKNVNDKNNSDXYIK 445

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters: -DEV-xlh
-MODEL-frame+pn.model -DB=xlh
-Q/cgn2.1/USPTO_SPOOL/US09890220/runat_13062003_144850_16438/app_query.fasta_1.583
-DB=N.Geneseq_101002 -OFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220.ecgn1.1.263.ernat_13062003_144850_16438 -NCPU=6 -ICPU=3
-NO_MAP -LARGEUDERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGCLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FCAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N.Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2378	100.0	1722	21	AAA63670	Nucleotide sequence
2	2378	100.0	1722	21	AAA47751	VNR2 nucleic acid.
3	2285	96.1	1715	21	AAA63671	Nucleotide sequenc
4	2285	96.1	1715	21	AAA47753	VNR2 nucleic acid.
5	2258.5	95.0	1737	21	AAA47755	Aberrantly spliced
6	1923	80.9	1497	21	AAC51581	Arabidopsis thailia
7	1813.5	76.3	6338	21	AAA47752	VNR2 nucleic acid.
8	1659	69.8	5895	21	AAA47754	VNR2 nucleic acid.
9	923	38.8	2280	21	AAZ36947	CDNA encoding a MP
10	906.5	38.1	2248	21	AAZ36953	CDNA encoding an O
11	451.5	19.0	558	21	AAA47758	At Hyp 2245035 (AT
12	407.5	17.1	5580	21	AAZ36948	DNA encoding a MPC
13	368.5	15.5	325	21	AAA47757	Arabidopsis thailia
14	340.5	14.3	640	21	AAC44147	Arabidopsis thailia
15	327	13.8	319	21	AAC51582	Arabidopsis thailia
16	324	13.6	525	21	AAC54919	Arabidopsis thailia
17	311	13.1	463	21	AAA47756	Rice expressed seq
18	225	9.5	3545	22	AAZ59565	Human CDNA encodin
19	221	9.3	4441	21	AAA47759	KIRA0160 CDNA. Ho
20	221	9.3	4441	24	ABL69961	Pancraas cancer re
21	219	9.2	3572	23	ABL29637	Drosophila melanog
22	208	8.7	7021	23	ABL29636	Drosophila melanog
23	196.5	8.3	1216	21	AAZ36962	DNA fragment encod
24	157.5	6.6	1257	24	AAZ62641	CDNA sequence #428
25	153	6.4	2481	22	AAZ5945	Human CDNA encodin
26	131.5	5.5	3738	21	AAA70178	Plasmodium falcipa
27	128.5	5.4	5454	21	AAA70236	Plasmodium falcipa
28	127.5	5.4	3300	21	AAZ56599	DNA encoding yeast
29	127	5.3	2466	23	ABL27385	Drosophila melanog
30	127	5.3	8045	21	AAA65171	Neurospora crassa
31	126	5.3	646	24	ABQ46952	Oligonucleotide fo
32	126	5.3	646	24	ABQ46953	Oligonucleotide fo
33	124.5	5.2	2061	21	AAW70228	Plasmodium falcipa
34	124	5.2	557	24	ABQ43458	Oligonucleotide fo
35	124	5.2	557	24	ABQ43459	Oligonucleotide fo
36	124	5.2	5340	22	AAZ62689	P. falciparum egfs
37	123.5	5.2	6033	21	AAA70152	Plasmodium falcipa
38	122.5	5.2	4311	21	AAA70133	Plasmodium falcipa
39	122	5.1	5145	21	AAZ70209	Plasmodium falcipa
40	121.5	5.1	412	21	ABN81085	Shimp polynucleot
41	120.5	5.1	198	17	AAV17204	DNA-spanner oligon
42	120.5	5.1	198	17	AAV17205	DNA-spanner oligon
43	120	5.0	7143	21	AAA70250	Plasmodium falcipa
44	119.5	5.0	567	21	AAA29550	HIV codon altered
45	119	5.0	3282	24	ABV79865	Fungal ZBC gene se

ALIGNMENTS

RESULT 1
ID AAA63670 standard; CDNA; 1722 BP.

AC AAA63670;
XX
DT 04-DEC-2000 (first entry)

DE Nucleotide sequence of VNR2 sequence of Arabidopsis sp.

XX H51: one locus-FRIGIDA; FRI gene; flowering time; blotting;
KW flower initiation; stem elongation; flower production; VNR2; ss.

XX Arabidopsis sp.

XX WO200046358-A2.

XX 10-AUG-2000.

XX 25-JAN-2000; 2000MO-GB00197.
 PF 05-FEB-1999; 99GB-0002660.
 XX (PLAN-) PLANT BIOSCIENCE LTD.
 PA Johnson U, West J, Dean C;
 FI WPI: 2000-532899/48.
 DR
 XX New nucleic acid derived from the FRI locus of a plant, e.g.
 PT Arabidopsis, encoding a polypeptide capable of specifically altering
 PT the flowering time of a plant -
 XX
 PS Example 5; Page 53-54; 73pp; English.
 CC The present sequence represents a VRN2 sequence from variety Landsberg
 CC erecta. The specification describes a sequence which encodes a
 CC polypeptide capable of specifically altering the flowering time of a
 CC plant. The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
 CC of Arabidopsis. The FRI polynucleotide is used to transform plants, so
 CC that the flowering time of a plant is altered. This is used, for example,
 CC for plants in which the leaves or tubers are a commercial product, where
 CC it is desirable to avoid 'blotting' (initiation of flowers and stem
 CC elongation) at too early a stage. Conversely, it may be desirable to
 CC alter flowering under certain circumstances e.g. to vary flower
 CC production across the seasons.
 XX
 SQ Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;

Alignment Scores:

Pred. No.: 3,53e-243 Length: 1722
 Score: 2378.00 Matches: 445
 Percent Similarity: 100.008 Conservative: 0
 Best Local Similarity: 100.008 Mismatches: 0
 Query Match: 100.008 Indels: 0
 DB: 21 Gaps: 0

US-09-890-220-2 (1-445) x AAA63670 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaValSerSerProGluGluValIleSerThrAspGlu 20
 DB 231 ATGTGTAGCGAGATGTCGCGGAAATCTCACGAGAGATGATTCAACGATGAG 290
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 291 AATCTCTTGATATATGTAAACCTGTCGACATATATTAACATCTTCACCTCGCTCTTA 350
 QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 351 GGCACACCATCGTTCTTCCAAAGATCTTGAACCTTGAACCTTGAACCTTGAACCTTGA 410
 QY 61 LysSerArgSerThrGlyMetValIlePheAsnTyrLysAspCysAsnAspThrLeuGln 80
 DB 411 AAGTCAGAGATCTAGGAGATGAGTTCACACTTAAGATTTGAATTAACACTTACAG 470
 QY 81 LysThrGluValArgLysAspCysSerProPheCysSerMetLeuCysGlySerPhe 100
 DB 471 AAAACGTAACTTGGAGAGATGTTCTTCTTCCATTTTGTCTCATTTGTGTGAGTTC 530
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluLysLeu 120
 DB 531 AAGGGCTGCAATTCATTTGAAATTCATCTCAATTAATTAATTAATTAATTAATTAATTA 590
 QY 121 PheGluLeuTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 591 TTGCAAAATATCACAGACATTAATGTTCTGTAATAAATTAATTAATTAATTAATTAATTA 650
 QY 141 GluGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 DB 651 GAAAGAGATGATGACATTAATTTGAGCCCTCTCTCTGCTGCAAACTGTAAGGCG 710

QY 161 ArgGlnArgGlyValArgAsnAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
 DB 711 AGACAAAGAGGTGGAGAAATTAACACCGAGAGACTTAATTAATTAATTAATTAATTAATTA 770
 QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
 DB 771 TCACCAGATTAACTAAATGAGCAAGAAATGAATTCACCTACTTAATTAATTAATTAATTA 830
 QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyIlePheGluMetThrSerAsnIle 220
 DB 831 GGTTAGGATATCCGAGCAACAGAGCTTCTGCAAAATTAATTAATTAATTAATTAATTAAT 890
 QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 DB 891 CCACAGCCATAGCCCACTCTTCTGTGACGCTGTCTTAAGTTAATTAATTAATTAATTAAT 950
 QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 DB 951 GCTGTGTCCTCGTACTTAAGACAAAGAAATTAATTAATTAATTAATTAATTAATTAATTA 1010
 QY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 DB 1011 CACCTACTTCTTCANAAAGCCCAATGCTATCATTTCTCACAGAGTCCAGCCAAATGGCGCTT 1070
 QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
 DB 1071 GAGCAAGTATGTCGACCGGATAGCGAGATGACATGATGATGATGATGATGATGATGAT 1130
 QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 DB 1131 GAGAGATCGCAGATGCTGATGACTTGTGTGATGATGATGATGATGATGATGATGATGAT 1190
 QY 321 HisLeuTyrPasnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
 DB 1191 CATCTTTGAACTCGTTTGTAAGAAACAAAGGTTATACAGATGATGATGATGATGATGAT 1250
 QY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
 DB 1251 GCATGTGAAGCATTTTCAAGATTTTACAGAAAGATGTCACCGTACTCAACACTCTTC 1310
 QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
 DB 1311 TGGTGTGAGATGTTTGTGATTAACCTATGAGAACCATGATGATGATGATGATGATGAT 1370
 QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 DB 1371 ATCAACAACCTGCAATACCATCTCGAGAAATTCGCGTAATGCTCAACACACACACAC 1430
 QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleVal 420
 DB 1431 AACCAACAACAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1490
 QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
 DB 1491 GATCATCCCATGACATTAACAACAACAACAATGTTGACAAACAAGACATTAACAACGACA 1550
 QY 441 AspLysValIleLys 445
 DB 1551 GACCAAGTAAATTA 1565

RESULT 2

AAA47751 standard; cDNA: 1722 BP.

AAA47751;

16-NOV-2000 (first entry)

VRN2 nucleic acid.

VRN2 gene; VRN2; plant characteristic; flowering time;
 leaf size; leaf shape; shade avoidance response; reproduction;
 breeding; pollination; cultivation; ss.

```

XX Arabidopsis thaliana var 'Landsberg erecta'
OS
XX Key Location/Qualifiers
FH 231..1568
FT CDS
FT mutation
FT /tag= a
FT /product= "VRN2 polypeptide"
FT replace(1199,A)
FT /tag= b
FT /note= "found in mutant vrn2-1; alters a Trp codon
FT to a stop codon leading to a truncation of
FT the protein"
XX
XX MO200044918-A1.
XX
XX PD 03-AUG-2000.
XX
XX PF 28-JAN-2000; 2000WO-GB00248.
XX
XX PR 28-JAN-1999; 99GB-0001927.
XX
XX (PLAN-) PLANT BIOSCIENCE LTD.
XX
XX PI Dean C, Gendall A;
XX
XX DR WPI: 2000-499333/44.
XX P-PSDB: AAB00060.
XX
XX PT Isolated vernalization gene VRN2 is used to produce transgenic plants
XX with altered vernalization response, flowering time, leaf size and/or
XX shape or shade avoidance response for maximized reproductive success
XX
XX PS Claim 4, Fig 6; 105pp; English.
XX
XX CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
XX plant encode polypeptides which are capable of affecting one or more
XX vernalization responses such as, flowering time, leaf size and/or
XX shape or the shade avoidance response of a plant into which the
XX nucleic acid is introduced. Introducing such sequences into plants
XX to alter these characteristics maximises the reproductive success of
XX the plant.
XX
XX SQ Sequence 1722 BP; 538 A; 339 C; 348 G; 497 T; 0 other;
XX
Alignment Scores:
Pred. No.: 3 53e-243 Length: 1722
Score: 2378.00 Matches: 445
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-890-220-2 (1-445) x AAA47751 (1-1722)
QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 231 AGGTGTAGGCAAAATGTCGCGCGAAATCTCACGAGAGATGATTCAACTGATGAG 290
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 291 AATCTCTGTATATATGTAAACCTGTCACATATATAAACATCTTCCACCTTGCTCTCTA 350
QY 41 GluAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 351 GGCACACCATCTTCTTCTTCCAGATGCTTGAACCTCAAAATGAGCAAGCGCAAAAGA 410
QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 411 AAGTCAAGATCTACTGGGATGTGTAGTTTCAACTATTAAGATGTGTATTAACACATTACG 470
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 471 AAAACTGAAGTTAGGAGATGTCTTCTCATTTTCTCTATGCTATGTGTAGTCTTTC 530

```

```

QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 531 AAGGGGCTGCAATTCATTTGATTCATCTCATGATTATTTGAATTAAGTCAAGCTT 590
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 591 TTCGAAGATATACAGACACTTAATGTTCTGTAAAACTTAATTCCTCATATTGAGGAA 650
QY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 651 GAAGGAAGTGAAGACATTAATTTGAGCCCTCTCTCTCTCTCTCTCAAAACCTGTAAAGCG 710
QY 161 ArgGlnArgGlyAlaArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB 711 AGACAAAGAGGTGGGCAAAATACACCGAGACACTTAAGATGCTTTTACCGTTGGAT 770
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArg 200
DB 771 TCACCCAGTTTAACATAATGSCACAGAAATGGAATCACCTTAATGATGAAACCGT 830
QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 831 GGTTAGGATATCCCGAGCGCAACAGACTTCTGCAATTTGAGATGACCAACACTT 890
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 891 CCACCAAGCCATAGCCCACTCTCTCTGAGCGCTGGGTAAAGTTATATTGACAAACGGA 950
QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB 951 GCTGTGGTCCCTGCTACTAGACAAAGAAAGTTATCTGCTGAGCGATCAGAGCTTAAGC 1010
QY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1011 CACCTTAATCTTCAGAAAGCCCAATTCATCTTCACAGAGTCCAGCAATGGCGCTT 1070
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAspPhe 300
DB 1071 GAGCAAGTATATGTCTGACCGGGATGACCGAGATGAAGTGAATGATTCACATGTTT 1130
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB 1131 GAAGATCGCCAGATGCTTGATGATCTTGTGATGTGAATGAATGAAGAAACCAATTG 1190
QY 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB 1191 CATCTTTGGAACTCGTTGTAGAAACAAAGGTTATGACAGATGCTATATCTCTTGG 1250
QY 341 AlaCysGluAlaPheSerArgPheTyrGlyLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1251 GCATGTGAAGCATTTTCAAGATTTTACGAAAGAGTTCCACCGTTACTCATCACTCTTC 1310
QY 361 TrpCysTrpArgLeuPheLeuIleLysLeuThrPasnHisGlyLeuValAspSerAlaThr 380
DB 1311 TGGTGTGGAGATTTGTTTGAATTAACATATGAACATGAGACTTGTGACGACGACCC 1370
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
DB 1371 ATCAACAACTGCAATACATCTCTCGAATTTCCCGTAAATGCTCAACACCCACACACC 1430
QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
DB 1431 AACCAACACACAGTGTGATGTCTCCAGTACTCAACACCAACACAAATTAACATTG 1490
QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
DB 1491 GATCATCCCAATGACATAAACAACAAGAAATGTTGACAAACAGACAAATTAACGACA 1550
QY 441 AspLysValIleLys 445
DB 1551 GACAAAGTAATTA 1565

```

```

RESULT 3
AAA63671
ID AAA63671 standard; cDNA; 1715 BP.
XX
AC AAA63671;
XX
DF 04-DEC-2000 (first entry)
XX
DE Nucleotide sequence of VRN2 sequence of Arabidopsis sp.
XX
KW H51: one locus-FRIGIDA; FRI gene; flowering time; blotting;
XX flower initiation; stem elongation; flower production; VRN2; ss.
OS Arabidopsis sp.
PN WO200046358-A2.
PD 10-AUG-2000.
PE 25-JAN-2000; 2000MO-GH00197.
PR 05-FEB-1999; 99GB-0002660.
PA (PLAN-) PLANT BIOSCIENCE LTD.
PI Johanson U, West J, Dean C;
XX WPI: 2000-532899/48.
XX
XX New nucleic acid derived from the FRI locus of a plant, e.g.
XX Arabidopsis, encoding a polypeptide capable of specifically altering
XX the flowering time of a plant.
XX
XX Example 5: Page 54-55; 73pp; English.
XX
XX The present sequence represents a VRN2 sequence from variety Columbia.
XX The specification describes a sequence which encodes a polypeptide
XX capable of specifically altering the flowering time of a plant.
XX The polypeptide is encoded by the FRI (one locus-FRIGIDA) locus
XX of Arabidopsis. The FRI polynucleotide is used to transform plants, so
XX that the flowering time of a plant is altered. This is used, for example,
XX for plants in which the leaves or tubers are a commercial product, where
XX it is desirable to avoid 'blotting' (initiation of flowers and stem
XX elongation) at too early a stage. Conversely, it may be desirable to
XX alter flowering under certain circumstances e.g. to vary flower
XX production across the seasons.
XX
XX
XX Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;
XX
Alignment Scores:
Pred. No.: 2,88e-233 Length: 1715
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 21 Gaps: 1
DB:
US-09-890-220-2 (1-445) x AAA63671 (1-1715)
QY 1 MetCysArgGlnAsnGlySerProGluGluValIleSerThrAspGlu 20
DB 233 ATGTGTAAGGAGATGTCGCGGAATCTTCACGAGAGATGATTTCACGTGATGAG 292
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 293 AATCTCTTGATATATGTAAACCTGTCACATATATACATCTTTCACCTTCGCTCTCTA 352
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 353 GGCACACCATCGCTTTCGCAAGATGCTTGAACTCAAAATTTGGGCAAAACGCAAAAGA 412
QY 61 LysSerArgSerThrIleMetValIlePheAsnTyrLysAspCysAsnAspThrLeuGln 80

```

```

DB 413 AAGTCACAGATCTACTGGGATGGTGTGTTTCAACTATATAGCATTTGTAATATCATTTACA 472
QY 81 LysThrGluValArgLysAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
DB 473 AGAACTGAAGTAAAGGAGATGTTGTTGTCATTTGCTATTTGATGATGATGATGATGATG 532
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlyLeu 120
DB 533 AAGGGCTGCAATTTTCATTTGATATTCATCTCATGATTTATTTAAATTTAGTTCAAGCTT 592
QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
DB 593 TTGGAAGAATACACAGACGTTAAAGTTTCTGTAATAACTTAACTTCAATATTTGAGCAA 652
QY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
DB 653 GAAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 712
QY 161 ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
DB 713 AGACAAAGAGGTGGGAGAAATATACACAGAGACTTAAGTATGCTTTTACCGTTGAT 772
QY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIlePheLeuAsnAspGlyLysArg 200
DB 773 TCACCCAGTTTACGTAATGTCAGCAAGAAATGGAATTCCTGCTGAATATGTAACCT 832
QY 201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB 833 GGTTAGAGATATCCCGAGGACAGACCTTGCGACAAATTTGAGATGACTGACAACTT 892
QY 221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB 893 CCACGACGATAGCCCACTCTTCTGTCAGCTGGCTGCAAAATGTAATTAACAACCGAA 952
QY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerLysIleGluArgSerGluAlaArgSer 260
DB 953 GCTGTGGTCCCTGCTACTAAGACAAAGAACTTATCTGTCAGGACAAAGCTGTGAAGC 1012
QY 261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB 1013 CACCTACTCTTCCAGAAAGCCCAATTCATCATCTTCACAGATCCAGCAATGCGCTT 1072
QY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
DB 1073 GAGCAAGTAATGTCGATCGGATGAGGAGATGAGATGATGATGATGATGATGATGAT 1132
QY 301 GluAspArgGlnMetLeuAspAspPheValAspValAspLysAspGluLysGlnPheMet 320
DB 1133 GAAAGATCCGCAAGATGCTTGAATGACTTGTGATGTGATGATGATGATGATGATGAT 1192
QY 321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerThr 340
DB 1193 CATCTTTGGAAGCTGCTTTGTAGAAACAAAGGTTTATGACAGATGCTATATCTCTTGG 1252
QY 341 AlaCysGluAlaPheSerArgPheTyrGlyGluLysGluLeuHisArgTyrSerSerLeuPhe 360
DB 1253 GCATGTGAGATTTTTCAGATTTTACGAGAAAGATGACAGTGTACATCACTCTTTC 1312
QY 361 TrpCysThrArgLeuPheLeuIleLysLeuThrTrpAsnHisGlyLeuValAspSerAlaThr 380
DB 1313 TGGTGTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1372
QY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrThr 400
DB 1373 ATCAACAACATCAATACCATCTTCGAGAAATTTGCCGTAATAACCTCA-----GTCACTAAC 1426
QY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAspAsnAsnAsnIleVal 420
DB 1427 AACCAACAACAACAGTGTGATCATCCAGAGTCAACAAACCAACAACAATATATTTGTG 1486
QY 421 AspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB 1487 GATCATCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1546

```


OY 441 Asplys 442
 DB 1547 GNCAG 1552
 RESULT 4
 AAA47753
 ID AAA47753 standard; cDNA; 1715 BP.
 XX
 AC AAA47753;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE VRN2 nucleic acid.
 XX
 KW Vernalization gene; VRN2; plant characteristic; flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ss.
 OS Arabidopsis thaliana var 'Columbia'.
 XX
 PN WO200044918-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-GB00248.
 XX
 PR 28-JAN-1999; 99GB-0001927.
 XX
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 XX
 PI Dean C. Gendall A;
 XX
 DR WPI: 2000-499333/44.
 DR P-PSDB: AAB00061.
 XX
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 XX
 PS Claim 5; Page 71; 105pp; English.
 XX
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.
 XX
 SQ Sequence 1715 BP; 536 A; 326 C; 353 G; 500 T; 0 other;
 Alignment Scores:
 Pred. No.: 2,886-233 Length: 1715
 Score: 2285.00 Matches: 428
 Percent Similarity: 97.51% Conservativeness: 3
 Best Local Similarity: 96.83% Mismatches: 9
 Query Match: 96.09% Indels: 2
 DB: 21 Gaps: 1
 US-09-890-220-2 (1-445) x AAA47753 (1-1715)
 OY 1 MetCysArgGlnAsnGlySerProGluValIleSerThrAspGlu 20
 DB 233 ATGTATAGGCAAAATGTCGCGAAATCTCACCAGGAGATGATTTCAACTGATGAG 292
 OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 299 AATCTCTGATATATGTAACCTGTGACATAATAACATCTTTCACCTTGCTCTCTCA 352
 OY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 353 GGCAACCCATCGTTCTCTGCCAAGATGCTTGAACACTACAAAATTGGGCGCAAGCGCAAAAGA 412

OY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnThrLeuGln 80
 DB 413 AAGTCAAGATCTACTGGGATGTGCTTTCACACTATAGGATTTGATATATACATTACAA 472
 OY 81 LysThrGluValArgLysPheCysSerCysProPheCysSerMetLeuCysGlySerPhe 100
 DB 473 AGAAGTGAAGTGAAGGAGATGTCTGTGTCATTTTGTCTATAGCTATGTGTGCTTC 532
 OY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 DB 533 AAGGCGCTGCATTTTCATTTGAAATTCATCTCATGATTTATTTGAAATTTGAGTCAAGCTT 592
 OY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 593 TTGGAAGATACCAACAGTTATATCTTGTGAACCTTAATCTTCAATTTTAGGA 652
 OY 141 GluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 DB 653 GAAGGAAGTGAATGATTAATTTGAGCCCTCTCTCTGCTGCGAAGCTGTAAAGCT 712
 OY 161 ArgGlnArgGlyLysArgAsnThrArgArgLeuLysValLysPheLeuProLeuAsp 180
 DB 713 AGACAAAGAGGTGGCAGAAATACACACAGAGACTTAAGATGCTTTTACCGTTGAT 772
 OY 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
 DB 773 TCACCCAGTTTACGTCATATGCGACAGAAATGGAATTCGCCCTGCTGATATGGAACCT 832
 OY 201 GlyLeuGlyTyrProGluValThrGluLeuValGlyGlnPheGluMetThrSerAsnIle 220
 DB 833 GGTTTAGATATCCGAGGCAACAGAGCTGTGTCGATTTGAGATGCTAGCAACAT 892
 OY 221 ProProIleIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 DB 893 CCACCAAGCCATGCCACACTCTCTGCGAGCTGTGAAGATTAATTAACACCGAA 952
 OY 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 DB 953 GCTGTGTCCTGCTACTAGACAAAGAAATATCTGCTGACGATCAGAGCTAGAAC 1012
 OY 261 HisLeuLeuGlnGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
 DB 1013 CACCTACTCTTCAGAAAGCCAAATCTTCAATTCACAGAGTCCAGCAATGGGGCTT 1072
 OY 281 GluGlnValMetSerAspArgAspSerGluAspGluValAlaAspAspValAlaAspPhe 300
 DB 1073 GAGCAAGTATGCTGATCGGATGCGAGATGAAGTCGATGACGATGTCAGATTTT 1132
 OY 301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
 DB 1133 GAAGATCGCCAGATGCTTATGATGATCTTGTGATGTGAATTAAGATGAAGCAATTCATG 1192
 OY 321 HisLeuTyrPasnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTyr 340
 DB 1193 CATCTTTGGAACCTCGTTTGAAGAAACAAAGGTTATAGAGATGGCATATCTCTGG 1252
 OY 341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
 DB 1253 GCATGGAAGTATTTCAAGATTTTACGGAAGAGATTCACCTGTACACACACTCTTC 1312
 OY 361 TrpCysTyrArgLeuPheLeuLeuLysLeuTyrPasnHisGlyLeuValAspSerAlaThr 380
 DB 1313 TGGTGTGGAGATTTGTTTGTATTAACATATGAACCATGAGCTTTCACATCAGCCACC 1372
 OY 381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
 DB 1373 ATCAACAACTGAATATCATCTCGAGATTCGCCGTAATAATCACTCA-----GTCAATAC 1426
 OY 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
 DB 1427 AACACACAAACAAAGTGTGATCATCCAGTCACTAAACACCAACAAACAAATTAACATGTG 1486

QY 421 Asph1SPROASnAsp11eAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
 Db 1487 GATCATCCGATGACATATAAAGAAAGAAACAAATGTTGACACACAGACATATACAGCAGA 1546
 QY 441 AspLys 442
 Db 1547 GACAAAG 1552

RESULT 5

AAAA7755
 ID AAA47755 standard; cDNA; 1737 BP.

AC AAA47755;

DT 16-NOV-2000 (first entry)

DE Aberrantly spliced VRN2 nucleic acid.

KM Vernalization gene: VRN2; plant characteristic: flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 XX breeding; pollination; cultivation; ss.

OS Arabidopsis thaliana var 'Columbia'

PN W0200044918-A1.

PD 03-AUG-2000.

PE 28-JAN-2000; 2000WO-GB00248.

XX 28-JAN-1999; 99GB-0001927.

PR (PLAN-) PLANT BIOSCIENCE LTD.

PA Dean C. Gendall A;

XX WPI: 2000-498333/44.

DR P-PSDB; AAB00062.

XX Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 XX
 PS Claim 8; Page 74-75; 105pp; English.

CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant. This cDNA was generated from an aberrantly spliced VRN2
 CC gene.

SO Sequence 1737 BP; 543 A; 330 C; 359 G; 505 T; 0 other;

Alignment Scores:

Pred. No.: 1,966-230 Length: 1737
 Score: 2258.50 Matches: 428
 Percent Similarity: 95.78% Conservative: 3
 Best Local Similarity: 95.11% Mismatches: 9
 Query Match: 94.97% Indels: 10
 DB: 21 Gaps: 2

US-09-890-220-2 (1-445) x AAA47755 (1-1737)

QY 1 MetCysArgGlnAsnGlySerValAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 Db 233 ATGTGTAGGAGAGATGTCGCGGAAATCTCACCGGAGAGATTTTCAACGATGAG 292
 QY 21 AsnLeuLeuIleTyrCysLysProValAlaArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 Db 293 AATCTCTGTATATATGTAAACCGTGTGACATATATAACATCTTCACTTCGCTCTCTA 352

QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrIleGlyAlaLysArgLysArg 60
 Db 353 GGCAACCATCGCTTTCTGCCAAGATCGTTGAACTACAAAATGCGGCAAGCGCAAAAGA 412
 QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
 Db 413 AAGTCAAGATCTACTCGGATGATGTTTCACTATAGATGATTTATATACATTACAA 472
 QY 81 LysThrGlnValAlaArgLysAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 Db 473 AGAACTGAAGTTAGGAGAGATTTCTTGTGCATTTTGCCTATGCTATGCTGACCTTC 532
 QY 101 Lys-----GlyLeuGlnPheHisLeuAsnSerHisAspLe 113
 Db 533 AAGGTGGCAACTATTACAACTGAGGGCTGCATTTTCAATTGATTCATCTCATGATTT 592
 QY 113 nPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLysLe 133
 Db 593 ATTTGAATTTGATTCAAGCTTTTGAGAGATACACAGATTAATGTTCTGTATAACT 652
 QY 133 uAsnSerPheIlePheGluGluGluGlySerAspAspLysPheGluProPheSerLe 153
 Db 653 TAATTCCTTCATATTTGAGAGAGAGATGATGATTAATTTGAGCCCTCTCTCT 712
 QY 153 uCysSerLysProArgLysArgArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLys 173
 Db 713 CTGCTCGAAACCTCGTAGCGGTAGCAAAAGAGTGGCAGAAATMACACGAGAGACTTAA 772
 QY 173 sValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyLeuTh 193
 Db 773 AGTATGCTTTTACCCTGTTGATCCACCACTTTAGATGGCACAGAAATGGAATTC 832
 QY 193 lLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyL 213
 Db 833 CCTGCTGAATATGGAACCCGTGTTTATGATATCCCGAGGCAACGACCTTGCCTGCA 892
 QY 213 nPheGluMetThrSerAsnIleProPheAlaIleAlaHisSerSerLeuAspAlaGlyAl 233
 Db 893 ATTTGAGATGACTGACAAACCTCCACGACCATATGCCCCTCTCTGAGCGCTGAGC 952
 QY 233 aLysValIleLeuThrSerGluAlaValAlaProAlaThrLysThrArgLysLeuSerAl 253
 Db 953 TAAAGTTATATTAAACAACCGAAGCGTGTGCTCCTACTAAGACAAAGTATATCTGC 1012
 QY 253 aGluArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
 Db 1013 TGAGCGATCAGAGCGCTGAAAGCCACTACTTTCAGAAACGCCAATTCTATCATCTTCA 1072
 QY 273 sArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluVa 293
 Db 1073 CAGAGTCCAGCAATGGCGCTTGAGCAATGATGTCTGATCGGAGATAGGAGATGAAGT 1132
 QY 293 lAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAs 313
 Db 1133 CGATGACGATTTGACAGATTTTGAAGATCCCGAGATCGCTTGTGATGATCTGATGATAA 1192
 QY 313 nLysAspGluLysGlnPheMetHisLeuThrPasnSerPheValArgLysGlnArgValAl 333
 Db 1193 TTAAGATTAAGAAAGATCATGCTTGTGAACTGCTTGTGTAAGAAACAAAGGCTAT 1252
 QY 333 eAlaAspGlyHisIleSerThrPalacysgLysGluAlaPheSerArgPheTyrGluLysGluLe 353
 Db 1253 AGCAGATGCTATATTCCTTGGCGATGCTGAAGATTTTCAAAATTTTTCGAGAAAGACTT 1312
 QY 353 uHisArgTyrSerSerLeuPheThrPcysThrPargLeuPheLeuIleLysLeuThrPasnH 373
 Db 1313 GCACGTGTACTACATCACTCTTCTGCTGAGATATGTTTGAATTAACATATGAAACCA 1372
 QY 373 sGlyLeuValAspSerAlaThrIleAsnAsnGlyAsnThrIleLeuGluAsnCysArgAs 393
 Db 1373 TGACATGCTGACTCAGCCACCATCAAAACATCATACATCTCTCGAAGATTCGCGTAA 1432

Oy		393	nSerSeraSpHrThrThrAsnAsnAsnAsnSeryAlasParpProSerASpSeras	413
Dd		1433	TACCTCA-----GTCACTAACCAACACMACAGTGTGATCATCCCGACTGCATAA	1486
Oy		413	nThraSnasnsnaSnIeValasPHisProsnAspIleasnInlysaSnasnnValas	433
Dd		1487	CACCACACACATATACATTGGATCATCCGAATGACATTAAMAAACAAGAACAATGTGA	1546
Oy		433	PasnllySnasnasnsnSerArGasPlys	442
Dd		1547	CAACACAGCACATTAACACAGCAGACGAAG	1574
RESULT 6				
AACS1581				
ID			AACS1581 standard; DNA; 1497 BP.	
XX				
AC			AACS1581;	
XX				
DT			18-OCT-2000 (first entry)	
XX				
XX			Arabidopsis thaliana DNA fragment SEQ ID NO: 69038.	
DE				
XX				
KW			Hybridisation assay; genetic mapping; gene expression control;	
KM			protein identification; signal transduction pathway;	
KW			metabolic pathway; promoter; termination sequence; ss.	
XX				
OS			Arabidopsis thaliana.	
XX				
PN			EP1033405-A2.	
XX				
PD			06-SEP-2000.	
XX				
PF			25-FEB-2000; 2000EP-0301439.	
XX				
ER		25-FEB-1999;	99US-0121825.	
PR		05-MAR-1999;	99US-0123180.	
PR		09-MAR-1999;	99US-0123548.	
PR		23-MAR-1999;	99US-0125788.	
PR		25-MAR-1999;	99US-0126264.	
PR		29-MAR-1999;	99US-0127685.	
PR		01-APR-1999;	99US-0127462.	
PR		06-APR-1999;	99US-0128234.	
PR		08-APR-1999;	99US-0128714.	
PR		16-APR-1999;	99US-0129845.	
PR		19-APR-1999;	99US-0130077.	
PR		21-APR-1999;	99US-0130449.	
PR		23-APR-1999;	99US-0130510.	
PR		23-APR-1999;	99US-0130891.	
PR		28-APR-1999;	99US-0131449.	
PR		30-APR-1999;	99US-0132408.	
PR		30-APR-1999;	99US-0132407.	
PR		04-MAY-1999;	99US-0132484.	
PR		05-MAY-1999;	99US-0132485.	
PR		06-MAY-1999;	99US-0132486.	
PR		06-MAY-1999;	99US-0132487.	
PR		07-MAY-1999;	99US-0132863.	
PR		11-MAY-1999;	99US-0134256.	
PR		14-MAY-1999;	99US-0134218.	
PR		14-MAY-1999;	99US-0134219.	
PR		14-MAY-1999;	99US-0134221.	
PR		14-MAY-1999;	99US-0134370.	
PR		18-MAY-1999;	99US-0134768.	
PR		19-MAY-1999;	99US-0134941.	
PR		20-MAY-1999;	99US-0135124.	
PR		21-MAY-1999;	99US-0135353.	
PR		24-MAY-1999;	99US-0135629.	
PR		25-MAY-1999;	99US-0136021.	
PR		27-MAY-1999;	99US-0136392.	
PR		28-MAY-1999;	99US-0136782.	
PR		01-JUN-1999;	99US-0137222.	
PR		03-JUN-1999;	99US-0137528.	
PR		04-JUN-1999;	99US-0137502.	
PR		07-JUN-1999;	99US-0137724.	

PR	08-JUN-1999;	9905-0138094
PR	10-JUN-1999;	9905-0138540
PR	10-JUN-1999;	9905-0138847
PR	14-JUN-1999;	9905-0139141
PR	16-JUN-1999;	9905-0139452
PR	16-JUN-1999;	9905-0139453
PR	17-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139454
PR	18-JUN-1999;	9905-0139455
PR	18-JUN-1999;	9905-0139456
PR	18-JUN-1999;	9905-0139457
PR	18-JUN-1999;	9905-0139458
PR	18-JUN-1999;	9905-0139459
PR	18-JUN-1999;	9905-0139460
PR	18-JUN-1999;	9905-0139461
PR	21-JUN-1999;	9905-0139817
PR	22-JUN-1999;	9905-0140353
PR	23-JUN-1999;	9905-0140354
PR	23-JUN-1999;	9905-0140354
PR	24-JUN-1999;	9905-0140635
PR	28-JUN-1999;	9905-0140823
PR	30-JUN-1999;	9905-0140991
PR	30-JUN-1999;	9905-0141287
PR	01-JUL-1999;	9905-0141842
PR	01-JUL-1999;	9905-0142154
PR	02-JUL-1999;	9905-0142055
PR	06-JUL-1999;	9905-0142390
PR	08-JUL-1999;	9905-0142803
PR	09-JUL-1999;	9905-0142920
PR	12-JUL-1999;	9905-0142977
PR	13-JUL-1999;	9905-0143542
PR	14-JUL-1999;	9905-0143624
PR	15-JUL-1999;	9905-0144005
PR	16-JUL-1999;	9905-0144085
PR	16-JUL-1999;	9905-0144086
PR	19-JUL-1999;	9905-0144332
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144331
PR	19-JUL-1999;	9905-0144333
PR	19-JUL-1999;	9905-0144334
PR	19-JUL-1999;	9905-0144335
PR	20-JUL-1999;	9905-0144352
PR	20-JUL-1999;	9905-0144632
PR	20-JUL-1999;	9905-0144884
PR	21-JUL-1999;	9905-0144814
PR	21-JUL-1999;	9905-0145086
PR	21-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145085
PR	22-JUL-1999;	9905-0145087
PR	22-JUL-1999;	9905-0145089
PR	22-JUL-1999;	9905-0145192
PR	23-JUL-1999;	9905-0145145
PR	23-JUL-1999;	9905-0145218
PR	23-JUL-1999;	9905-0145224
PR	26-JUL-1999;	9905-0145276
PR	27-JUL-1999;	9905-0145913
PR	27-JUL-1999;	9905-0145918
PR	27-JUL-1999;	9905-0145919
PR	28-JUL-1999;	9905-0145951
PR	02-AUG-1999;	9905-0146386
PR	02-AUG-1999;	9905-0146388
PR	02-AUG-1999;	9905-0146389
PR	03-AUG-1999;	9905-0147038
PR	04-AUG-1999;	9905-0147204
PR	04-AUG-1999;	9905-0147302
PR	05-AUG-1999;	9905-0147192
PR	05-AUG-1999;	9905-0147260
PR	06-AUG-1999;	9905-0147303
PR	06-AUG-1999;	9905-0147416

PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149930.
PR 23-AUG-1999; 99US-0149932.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155186.
PR 24-SEP-1999; 99US-0155659.
PR 26-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156599.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159337.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160810.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 9,16e-195
Score: 1923.00
Percent Similarity: 90.10%

Length: 1497
Matches: 370
Conservative: 3

Best Local Similarity:	89.37%	Mismatches:	7
Query Match:	80.87%	Indels:	34
DB:	21	Gaps:	3
US-09-890-220-2 (1-445) x AAC51581 (1-1497)			
QY 43 ProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysSer 62			
DB 322 CCATCGTTCTTCCAAAGTGGCTTGAATACAAAATTGGACCAAGCGCAAAAGAAATCA 381			
QY 63 ArgSerThrGlyMetValPheAsnTyrLysAspCysAsnAsnThrLeuGlnTyr 82			
DB 382 ACATCTACTGGGATGGATGTTTCACTATATAGGATTTGATATACACATTACGAAACT 441			
QY 83 GluValArgLysAspCysSerCysProPheCysSerMetLeuGlySerPheLysGly 102			
DB 442 GAACTTAGCGAGATGTTCTTCCATTTTGGCTATAGTATGTTAGCTTCAAGGG 501			
QY 103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 122			
DB 502 CTGCAATTCATTGCAATTCATTCATGATTTATTTGATTTGATTTGATTTGATTTG 552			
QY 123 GluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 142			
DB 553 ---TACATCTCTGTAATCCAAACCAATATGTTTGCAGC----- 588			
QY 143 SerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArgGln 162			
DB 589 -----TCGAACCTCTGTAAGCGGAGACAA 612			
QY 163 ArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182			
DB 613 AGAGGTGGCAAGAAATACACACAGAGACCTTAAAGTATGCTTTTACCGTTGATTCACC 672			
QY 183 SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspLysAsnArgLysLeu 202			
DB 673 AGTTTAACTAATGGCACAGAAATGGAATCCCTCTTAAAGATGGAACCCGTGTTTA 732			
QY 203 GlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPro 222			
DB 733 GGATATCCCGAGGACAGAACTTGGCGACATTTGAGATGACCGCAACCTTCCACCA 792			
QY 223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVal 242			
DB 793 GCATAGCCCACTCTCTCTGACGCGTGAAGTATATATGCAAGCAAGCAAGCTGTG 852			
QY 243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259			
DB 853 GTCCCTGCTACTAAGACAAAGATATATGCTGACGATCAGAGCTAGAAAGTACTGCC 912			
QY 260 -----SerHisLeuLeuGlnLysArgGlnPheTyrHis 271			
DB 913 ATATTAAACCGAAGAAATTTCTTCCAGGACACTCTCTTCAAGAAAGCCCAATTCATCAT 972			
QY 272 SerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArgAspSerGluAsp 291			
DB 973 TCTCACAGAGTCCAGCCCAATGCGCTTGACCAAGTATATCTGACCGGAGATACAGAGAT 1032			
QY 292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAsp 311			
DB 1033 GAACTCGATGACGATGTTGCGATTTTGAAGATCGCCAAATGCTTGAAGACTTGTGAT 1092			
QY 312 ValAsnLysAspGluLysGlnPheMetHisLeuTyrAsnSerPheValArgLysGlnArg 331			
DB 1093 GTGATATAAGATGAAAGCAATTCATGATCTTTGGAACTCTTTGTAAGAAACCAAGG 1152			
QY 332 ValIleAlaAspGlyHisIleSerTyrPalaCysGluAlaPheSerArgPheTyrGluLys 351			
DB 1153 GTTATACAGATGGTCAATATTTCTGCGCATGGAAGCATTTTCAAGATTTTACGAGAA 1212			
QY 352 GluLeuHisArgTyrSerSerLeuPheTyrPalaCysGluArgLeuLeuLysLeuTyr 371			
DB 1213 GAGTTGACCGGTTACTCATCATCTCTCTGTTGTGAGATTTTGTGATTAACATATAGG 1272			

QY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
 |||||||
 Db 1273 AACCATGGACTTGTGACTCACCCACCATCAACATGCAATACCATCTCCAGATTGG 1332
 QY 392 ArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAspArgProSerAsp 411
 |||||||
 Db 1333 CGTATAGCTGACGACACCCACCCACCAACACACAGTGTGATGTCCAGTGC 1392
 QY 412 SerAsnThrAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsn 431
 |||||||
 Db 1393 TCAAACACCAACAAACAAATACATTTGTGATCATCCCAATGACATTAACAAACAAACAAAT 1452
 QY 432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
 |||||||
 Db 1453 GTTGACAAACAGGACAAATACAGCAGACAAAGTAATTA 1494
 RESULT 7
 AAA47752
 ID AAA47752 standard; DNA; 6338 BP.
 XX
 AC AAA47752;
 XX
 DT 16-NOV-2000 (first entry)
 XX
 DE VRN2 nucleic acid.
 XX
 KM Vernalization gene; VRN2: plant characteristic; flowering time;
 KM leaf size; leaf shape; shade avoidance response; reproduction;
 KM breeding; pollination; cultivation; ds.
 XX
 OS Arabidopsis thaliana var 'Landsberg erecta'.
 XX
 PN WO200044918-A1.
 XX
 PD 03-AUG-2000.
 XX
 PF 28-JAN-2000; 2000MO-GB00248.
 XX
 PR 28-JAN-1999; 99GB-0001927.
 XX
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 XX
 PI Dean C, Gendall A;
 XX
 DR MPI; 2000-499333/44.
 XX
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 XX
 PS Claim 9; Page 68-70; 105pp; English.
 XX
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.
 XX
 SQ Sequence 6338 BP; 1932 A; 1076 C; 1134 G; 2193 T; 3 other;
 Alignment Scores:
 Pred. No.: 3 76e-182 Length: 6338
 Score: 1813.50 Matches: 440
 Percent Similarity: 43.58% Conservative: 1
 Best Local Similarity: 43.48% Mismatches: 4
 Query Match: 76.26% Indels: 569
 DB: 21 Gaps: 13
 US-09-890-220-2 (1-445) x AAA47752 (1-6338)

QY 1 MetCysArgGluAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu 20
 |||||||
 Db 2342 ATGTGAGGCACAAATTTGTGCGCCGAATCTCCACCGAGAGATGATTTCCACTGATGAG 2401
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 |||||||
 Db 2402 AATCTCTGATATATGTTAAACCTGTGACATATATACATCTTCACCTGCTGCTGCTA 2461
 QY 41 GlyAsn----- 42
 |||||||
 Db 2462 GCCAACGATGATTTGCCCTTCCTCTCATCATATTAGCTAGTAATCTTCATCTCTG 2521
 QY 42 ----- 42
 2522 TGTAGATCCACCACTAATAGTTGAGTTTGTAGCTGATATAGTCTGATTCATGCGCA 2581
 QY 43 -----ProSerPhe 45
 |||||||
 Db 2582 GTGTGCTTCTTTTGTCTCTAAATTTGACCTGTTGTGTTGTCAGCCACGCTT 2641
 QY 46 LeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer----- 62
 |||||||
 Db 2642 CTTCCAGATGCTTGAACTACAAAAATTGAGGAAAGCGCAAAAGAAAGTA-TGTTTCTT 2700
 QY 62 ----- 62
 2701 CTTGATGCTGCTGCTACAGTATAGTTATTTATCTTACTTAATGAAAGCTGATG 2760
 QY 62 ----- 62
 2761 ACCATATTTATCTTGTGTGATGATATGACATATGATGTTCTTCTTGTGTCAG 2820
 QY 63 -----ArgSerThrGlyMetVal 68
 |||||||
 Db 2821 CTATAACTTACATTTTATATAAATGTGTTTGGTTAGTCAGATCTACTGGAGTGA 2880
 QY 69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGluLysThrGlu----- 83
 |||||||
 Db 2881 GTTTCACTATAGATGTTGTAATACATATTACAGAAACGAAAGTATGCTTTTCT 2940
 QY 83 ----- 83
 2941 GTTCGTGACAAATTCGATGCAATGCTATGTTCTCTAGATGTTGTTATTACTA 3000
 QY 84 -----ValArgLysAspCysSerCysProPheCysSerMetIle 96
 |||||||
 Db 3001 TTTTTCGTATGTCATGACGATGAGGAGATGTTCTTGTCCATTTTGTCTATGCT 3060
 QY 96 uCysGlySerPheLys----- 101
 |||||||
 Db 3061 ATGTGTACTCTCAAGGTGGGCAACTATTACAACGTAGSTTCTTCCGGGCGCTTCATA 3120
 QY 101 ----- 101
 3121 TCTAACCTGTGAATGCTACTGCTGTTTCATGCTGTATACCTTACTGTTGCTTACAT 3180
 QY 101 ----- 101
 3181 ATTTTGTGTTGTTGTTGTTGTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3240
 QY 102 -----GlyLeuG 104
 |||||||
 Db 3241 GAGAAACATGTCAGATGAGACTTACAAACAAATGCTGTGCTGCTATGACGGGCTG 3300
 QY 104 InPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu----- 120
 |||||||
 Db 3301 AATTCATTTGAAATCATCATGATTTTGAATTTGAAAGTATGATGTGTTTGA 3360
 QY 120 ----- 120
 3361 TGGAAATCTTGATTTGCTATGCTTATTAATGAGTTATAGTTAAAAAGGCTCTTC 3420
 QY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136

Db 3421 CTATTGTAGCTTTGGAGAAATACAGACAGTTAAAGTTCTGTAAACCTTAATTCCTTC 3480
 QY 137 ILepheglu----- 139
 Db 3481 ATATTGAGGCACTTACTTTAACTGGTTAAATGGAAATCCAGTAGCTGTGAAAT 3540
 QY 140 -----GluGluGlySerAspAspLysPheG1 149
 Db 3541 TTTGTTATATTCATCCCTTATTTGACTAGAGAAAGAGATGATGAGATTAATTTGA 3600
 QY 149 uPheSerLeu----- 153
 Db 3601 GCCCTTCTCTCTGTGTAACCCCTCAGAACCCCTTCGATTAAATACCTTAATAGCAGTAAC 3660
 QY 154 -----CysSerLy 156
 Db 3661 TCCTTTCCTCTCTGTGTAACCTCTCTGTAATCCACCACTTAATGTTTTCAGCTCGAA 3720
 QY 156 sProArgLysArgGlnArgGlyLysArgAsnAsnThrArgArgLysValCysPh 176
 Db 3721 ACCTGTAAAGGGAGACAAAGAGGTGGCAGAAATATACACAGAGACTTAAAGTATGCTT 3780
 QY 176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAs 196
 Db 3781 TTTACCGTTGGATTACACCCAGTTTAACTAATGACAGAAATGGAAATCACCCTTACTTAA 3840
 QY 196 nAsp----- 197
 Db 3841 TGATGTAATAATCATATCTTCTGTGCTTCTGTGCTTGTGCTTACAACTTCATATTACAG 3900
 QY 197 ----- 197
 Db 3901 AAGAAGATACAAAGGCCCTGATTGTTAGTTTGTACTCTCTCCGCAATCTCTTGCGGA 3960
 QY 198 -----GlyAsnArgGlyLe 202
 Db 3961 GGGATTGTTACCGAAGCATGATGTACAAATTAATGACGATGCTACAGGAAACCGGGTGT 4020
 QY 202 uGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetTrpSerAsnIleProPr 222
 Db 4021 AGGATATTCGCGAGGCAACAGARCTCTGTGACAAATTTGAGATGACCAACATTCCTCCAC 4080
 QY 222 alaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVal 242
 Db 4081 AGCCATAGCCCACTTCTCTGTGACCGCTGTGCTAAAGTTATATTGACAAAGCAAGCTGT 4140
 QY 242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
 Db 4141 GGTCCTGTACTAAGCAAGAAAGTTATCTGTGACGATCAGAGGCTAGAGGTTTGT 4200
 QY 259 ----- 259
 Db 4201 TCATCATGACACCCCGCATCATTAATTAACATTCCGTTGTACAAATGTTCTTCTTAAT 4260
 QY 260 -----SerHisLeuLe 263
 Db 4261 ATGATATAAGTGTATTATAGTACGCCATATTAACCGAGAAATTTCTTCCAGCACTACT 4320
 QY 263 uLeuGlnLysArgGlnPheTyrHisSerHisArgVal----- 275
 Db 4321 TCTTCAGAAAGCCCAATTCATCATCTTCACAGAGTCCAGGTGATCCAAATTCCTTGACC 4380
 QY 275 ----- 275
 Db 4381 TACTTCTTAGCATTTTCTTAAATGCTCATGATGATATCTTATCAAGCATACTTGT 4440
 QY 275 ----- 275
 Db 4441 TTGTCTCATCAAAATTTGATTTGATGATGATGATATCAACGCAAAATAGTATGTCA 4500
 QY 276 -----Gln-ProMet 279
 ||| |||||

Db 4501 TGTGTCTCCGTTTATTATGCCACTAACCAAAAAATGCAATGTTTCTGTACAAACCAATGG 4560
 QY 279 lAleuGluAlaValMetSerAspArgAspSerGlnAspGluValAspAspValAla 299
 Db 4561 CGCTTGAGCAAGTAATGTCTAGCCGGGATGACGGAGATGAAGTCATGATGATGTGACG 4620
 QY 299 sPheGluAspArg----- 303
 Db 4621 ATTTGAGATGCCAGGTATTCATGATTTCTTTCGCTCATTTAAATAGACACAGA 4680
 QY 304 -----GlnMetLeu 306
 Db 4681 AATGTATATGATGTAACTGTAATGCGTTTGAACCTTAAAAAAAGCTGCGAGATGCTT 4740
 QY 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
 Db 4741 GATGACTTGTGTGATGTGAATAAAGATGAAGAAACAAATTCATGATCTTTGGAACCTGTT 4800
 QY 327 ValArgLysGlnArgVal----- 332
 Db 4801 GTAAAGAAACAAAGGTA-ACCTCTCTTACACATGACAGACACAAAAACCTTAG 4859
 QY 332 ----- 332
 Db 4860 TCTTACATTCACATACCTGTCTAAATGATTTTGTCTTATGACACTTGACTCAATTATGAT 4919
 QY 333 -----IleAlaAspGlyHisIleSerTrpAlaCysGluAlaPheSer 346
 Db 4920 TGTATGTTTCAGGGTATAGCAGATGTGATATTTCTTGGCATGTGACATTTTCA 4979
 QY 347 ArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrp----- 361
 Db 4980 AGATTTTACGAGAAAGATGTGCACCGTTACTCATCTCTTGTGTAATATATGATACAC 5039
 QY 361 ----- 361
 Db 5040 AATCATATACAGACATACCTACATCAATCTGTTTGTGTTTCTGAAAAAAATA 5099
 QY 362 -----CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAsp 377
 Db 5100 AATATTTCCAGGTGTGGAGATGTGTTTATTAATCAATGAGACATGATGCTTCAC 5159
 QY 378 SerAlaTrpIleAsnAsnAsnAsnSerValAspArgProSerAspSerAspTrpAsnAsn 417
 Db 5160 TCAGCCACCATCAACAACTGCAATACCATCTCGAGAAATTCGCGTATATGCTCAGACAC 5219
 QY 398 ThrThrThrAsnAsnAsnAsnSerValAspArgProSerAspSerAspTrpAsnAsn 417
 Db 5220 ACCACACCAACAAACAAACAGTGTGATCGTCCAGTACTCAAAACCAACAAACAT 5279
 QY 418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAsnLysAspAsn 437
 Db 5280 AACATTTGGGTCTATCCCAATGACATTAACAAAGAAAGATGTGTACAAAGACAT 5339
 QY 438 AsnSerArgAspLysValIleLys 445
 Db 5340 AACACGACAGACAAAGTAATTA 5363
 RESULT 8
 AAA47754
 ID AAA47754 standard; DNA: 5895 BP.
 AC AAA47754;
 XX
 AC
 DT 16-NOV-2000 (first entry)
 XX
 DE VRN2 nucleic acid.
 XX
 DE
 KW Vernalization gene: VRN2; plant characteristic: flowering time;
 KW leaf size; leaf shape; shade avoidance response; reproduction;
 KW breeding; pollination; cultivation; ds.
 XX
 OS Arabidopsis thaliana var 'Columbia'.

XX WO200044918-A1.
 XX 03-AUG-2000.
 PD 28-JAN-2000; 2000WO-GB00248.
 XX 28-JAN-1999; 99GB-0001927.
 XX (PLAN) PLANT BIOSCIENCE LTD.
 XX Dean C, Gendall A;
 XX WPI; 2000-499333/44.
 DR
 XX Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximizing reproductive success
 XX
 PS Claim 9; Page 72-74; 105pp; English.
 XX
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximises the reproductive success of
 CC the plant.
 XX
 SQ Sequence 5895 BP; 1742 A; 1027 C; 1070 G; 2056 T; 0 other;

Alignment Scores:
 Pred. No.: 9,986-166 Length: 5895
 Score: 1659.00 Matches: 424
 Percent Similarity: 42.25% Conservative: 4
 Best Local Similarity: 41.86% Mismatches: 12
 Query Match: 69.76% Indels: 577
 DB: 21 Gaps: 14

US-09-890-220-2 (1-445) x AAA47754 (1-5895)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 1905 ATGTCTAGGCAAAATTTGCGCGCAATCCTCACCGAGAGATTTCACTGATGAG 1964
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 1965 AATCTCTGTGATATATGTAACCTGTGACATATTAACATCTTTCACCTGCTCTTA 2024
 QY 41 GlysAsn----- 42
 DB 2025 GGCAA-CGTATGATTTGGCCTTCCTCTCATCATTTTATAGCTTAATCTTTCATCTCC 2083
 QY 42 ----- 42
 DB 2084 TGTGTAGATACCACTAATAGTTGAGTTGCTAAGCTGATTAATGCTGACTGATGCG 2143
 QY 43 -----Pro 43
 DB 2144 GAGTGTGCTCTTTGTCTCTCTAATGTTATTTGAACTTGTGTTGTGTTGAGGCCA 2203
 QY 44 SerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer--- 62
 DB 2204 TCGTTTTCGCCAAGATGCTTGAACATACAAATTTGGGCAAGCGCAAAAGAAAGTA-TGC 2262
 QY 62 ----- 62
 DB 2263 GTTCTCTCTGAATGATGTTGCCACAGTATGTTATTTATCTTACTCTAATATGAA 2322
 QY 62 ----- 62
 DB 2323 GCTGATGAACATATTTATCTTGTGTGATGATGATGACATTAATGAATGCTTCTCTTGG 2382

QY 63 -----ArgSerThrGly 66
 DB 2383 TTTCATGCTATACACTTATATTTTACAAAATTTGTCTTGTCTAGGTCAGATCTACTGCG 2442
 QY 67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 DB 2443 ATGTAGATTTCACACTATAGATGTGTATATATACATTACAAAGAACGAAAGTTAGTCT 2502
 QY 83 ----- 83
 DB 2503 TTTTCTGTCTTCGACAAAATTCAGATGCAATGCTATGTTCTCTAGATGATTTGTTAT 2562
 QY 84 -----ValArgGlnAspCysSerCysProPheCysSe 94
 DB 2563 TTACATATTTTTCCTGATATGTCACGACGATTTGGAGAGATTTCTTGTGCTATTTGCTC 2622
 QY 94 rMetLeuCysGlySerPheLys----- 101
 DB 2623 TATGCTATGTGTACTTCACAGTGGGCAACTATTACAACTGAGGTTCTTCCGGGGCCT 2682
 QY 101 ----- 101
 DB 2683 TTCAATATCTACACTGTGAAATGCTACTGCGCTTAAGCTATATACCTTTCACGTGTGG 2742
 QY 101 ----- 101
 DB 2743 TTACATATTTTGTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 2802
 QY 102 -----G 102
 DB 2803 TTATCTGAGAAAACATGTTCCAGTTCGAGCTTACATCCATCTGCTGTGTATGACAGG 2862
 QY 102 lyleGlnPheHisLeuAsnSerHisAspLeuPheGlnPheLysLeuPhe- 121
 DB 2863 GGCTCAATTTATTTGAATTCATCATGATTTATTTGAATTTAGTTCAAGATATG-T 2921
 QY 121 ----- 121
 DB 2922 GGTTTTATGAAATTTCTTGTGTTTGGCTATGCGGTTAGTAAAGGTTATAGTTAAAG 2981
 QY 122 -----GluGlyTyrGlnThrValAsnValSerValLysLeu 134
 DB 2982 GGTCTTCTCTATTTAGCTTTTGGAAATATACAGATTAATGTTCTCTGTAACACTTA 3041
 QY 134 snSerPheIlePheGlu----- 139
 DB 3042 ATTCCTCTCATATTTGAGGTCACTTAACTGTTAATTTGGGAAATCTTATAGCTG 3101
 QY 140 -----GluGlnGlySerAspAsp 146
 DB 3102 GTGAAATTTGGTTTATATTCATCCTTATTTGTACTAGAGAAAGAAAGATGATGAT 3161
 QY 147 LysPheGlnProPheSerLeuCys----- 154
 DB 3162 AATTTTGAAGCCCTCTCTCTCTG-GTAACTCTAGAACCCCTTGATTAATACCTTAATA 3220
 QY 154 ----- 154
 DB 3221 GCAGTAACTCTGCTTCTTCTGTCAGTACTTCTCTAATAATCAACCAATGTTTTCG 3280
 QY 155 -----SerLysProArgLysArgArgGlnArgGlyArgAsnAsnThrArgArgLeuLys 173
 DB 3281 AGCTGAACACCTCGTAAGGCTGACAAAGAGGTGCGAAGAAATTAACCCGAGACTTAA 3340
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGlnAsnGlyIleThr 193
 DB 3341 GTATGCTTTTACCGTGTGATTCACCACTTATAGTAATGCGAAGAAATGAAATTTGCC 3400
 QY 194 LeuLeuAsnAsp----- 197
 DB 3401 CTGCTGAATGATGTAATAATCAACATCTTCTGTGTGATGCTGTTGGCTTAGAAGCTTC 3460
 QY 197 ----- 197

Db	3461	ATTTTACAGAAAGAGATACAAATGACGCTGGATGTTAGTTTGTGACTTCCTCGCATC	3520
OY	198	-----G1ys	199
Db	3521	TTCTTGAGAGGTAAATGTTACCAAGACTGATGACAAAATTAAATGGACTGCTACAGAA	3580
OY	199		
Db	3581	CCGTGGTTTAGGATATCCCGAGCAACAGAGCTTGCTGGACAAATTTGAGATGACTACAA	3640
OY	219		
Db	3641	CATTCCACGAGCATAGCCCACTCTTCTCGAGCTGGTGGCTAAAGTTAATTAACAC	3700
OY	239		
Db	3701	CGAAGGTGGTCCCTGCTACTAGACAAAGAAATTATCTGTGAGCATCAGAGCTAG	3760
OY	259	-----	259
Db	3761	AAGGTTGTTCAATCATGACACCCCGCTCATATTAATACATACCTGTTGTACAAATGTT	3820
OY	260	-----Se	260
Db	3821	CTTCTATTATGATAGATAGTGTTCCTGACTAGCCATTAATTAACGAGAAAATTTCTTCAC	3880
OY	260		
Db	3881	CCACTACTCTTCCAGAAAACCCCAATTCATCTATCTCCACAGAGTCCAGATCCAACT	3940
OY	276	-----	276
Db	3941	TCCTCACCACACTCTTAGCAGATTTCTTAAATTGCTCATGATATCTATCAAGC	4000
OY	276	-----	276
Db	4001	ATACTGGTTTGTTCTCATCTAAATTTGTATTTTGGATCTGTATGTATCAGCAAAAA	4060
OY	276	-----	276
Db	4061	ATTATGTCATGTTGTCTCCGTTTATTCGCCACTAACCAAAAACCTGCATGTTCTTGTA	4120
OY	277	-----PrometalaIaIeugInIvalMetSerAspArgSerGluAspGluIvalAspAs	295
Db	4121	CMACCCAATGCCCTTGACCAAGTAATGTCTGATCGGAGTAAGCGAAGATGAAGTCAATGA	4180
OY	295		
Db	4181	CGAATGTCAGATTTTGAAGATCGCCAGATTCATCATGATTTCTTCTGCTCAATTAAG	4240
OY	303	-----	303
Db	4241	TAGGCAACAGAAAATGGTATACGATGTAACTTGCTAATGGCTTTGAAACTTAAAAAGC	4300
OY	304	-----GImetIeuaSpAspPheValIaSpValaSnIySaspGluIySgInPheMetHISLeuT	323
Db	4301	TGCAGATGCTTGATGACTTGTGGATGTGAATTAAGATGAAGAACCAATTCATGCATCTTT	4360
OY	323		
Db	4361	GGAACTGCTTGTGTAAGAAAACAAAGTAACACTCTTCTTACACTTGAACACACACAAA	4420
OY	330	-----	330
Db	4421	AGACCTTATGCTTACATTCACATACCTGTCTAATATGATTCGCTTATGAACTTGAAGCT	4480
OY	331	-----ArgValIlealIaSpGluHISISerTrralIaCysGlu	343
Db	4481	CAAAATTATGATGATGTTGGACGGGTTATAGCAGTGTCTATCTCTGGAGATGTGA	4540
OY	344	AlaPheSerArgPheArgIuIySgIuIeunHISArgIySerSerIeunPheTrp	361

Dd		4541	GATATTTCAGATTTTACGAGAAGAGTGCACTGTACTCATTCAATCATCTCTTGTAATAT	4600
Oy		361	-----	361
Dd		4601	AGTAGACCACAACATATACAGACATACATACATCAATTTTTGTTTCGTGA	4660
Oy		362	-----CSTPACGLuePhelLeuLysLeuTPASnHs61	374
Dd		4661	AAGAAAATAAATAAAATTCAGAGTGTGGAGATGTGTTTTGATTAACATATGGAACACATGG	4720
Oy		374	YleuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSe	394
Dd		4721	ACTGTGCAGCTACGCCACCATCCAAACACTGCATATACCATCTCGAATAATYGGCGTAATAC	4780
Oy		394	rSerAspTrhTrhTrhTrhAsnAsnAsnSerValAspArgProSerAspSerAsnth	414
Dd		4781	CTCA-----GTCACCTAACCAACCAACCAACAGTGTGCATCCAGTCACTCANACAC	4834
Oy		414	IAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAs	434
Dd		4835	CAACAAACAATAACATTGTGGATCATCCGAATGACATATAAAAAACAAGAACATGTGGCAA	4899
Oy		434	nLysAspAsnAsnSerArgAspLys 442	
Dd		4895	CAAGGACAAATAACGACGAGCACAG 4919	
RESULT 9				
AAZ36947				
ID	AAZ36947	standard; cDNA; 2280 BP.		
XX	AAZ36947:			
XX	13-MAR-2000	(first entry)		
DE	cDNA encoding a Mpc1 protein having flowering regulating activity.			
xx	Flowering regulating activity; Mpc1; flowering; germination;			
KW	super early flowering mutation; altered flowering time;			
KX	flowering regulating gene; food crop; vegetable; flowering inhibition;			
KW	productivity; ss.			
OS	Arabidopsis thaliana.			
PH	Key	Location/Qualifiers		
FT	CDS	310..2145		
FT		/tag= a		
FT		/product= "flowering regulating protein"		
PN	EP967278-A2.			
PD	29-DEC-1999.			
PF	28-JUN-1999;	99EP-0305077.		
PR	26-JUN-1998;	98JP-0180065.		
PR	24-JUN-1999;	99JP-0179043.		
PA	(MITA) MITSUI CHEM INC.			
PI	Yoshida N, Kato Y, Takahashi S, Yanai Y, Hirasaka J, Miwa T.			
DR	WPI: 2000-064612/06.			
DR	P-PsDB: AAI53932.			
XX	Novel DNA used to produce transgenic plants with altered floral			
PT	regulation which can have increased crop yields -			
PS	Claim 6; Page 16-21, 53pp. English.			
CC	The present sequence encodes a protein having a flowering regulating			
CC	activity, which is designated Mpc1. The genomic sequence is given			
CC	in AAZ36948. A rice Mpc1s also disclosed in the specification. The			
CC	rice and Arabidopsis cDNAs show significant homology with each other.			

CC A naturally occurring mutation of the Mpc1 gene eliminates normal
 CC flowering regulating ability of plants, and leads to flowering
 CC immediately after germination (super early flowering mutation). The
 CC Mpc1 polynucleotide sequence can be used to produce plants with altered
 CC flowering times in comparison with wild type plants, by enhancing
 CC or inhibiting the expression of the flowering regulating gene. Antisense
 CC polynucleotides can be used to reproduce the effects of the mutated Mpc1
 CC gene. This alteration can be used to increase the yield of food crops.
 CC Flowering inhibition of vegetables increases their productivity.

xx Sequence 2280 BP; 663 A; 445 C; 521 G; 651 T; 0 other;

Alignment Scores:

Pred. No.:	6,95e-88	Length:	2280
Score:	923.00	Matches:	216
Percent Similarity:	43.23%	Conservative:	52
Best Local Similarity:	34.84%	Mismatches:	108
Query Match:	38.81%	Indels:	244
DB:	21	Gaps:	9

US-09-890-220-2 (1-445) x AA636947 (1-2280)

```

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      310 ATGTGCGCATGAAAGCTCCGCTGCTGCTATTTGCGAAGAGAGAGAGATGCTGCTGAAG 369
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB      :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      370 AGCTTGGCTGCTATGCAAGCTTTGAACTCTCAATATTCATTCAACGCCGCTGCTAT 429
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      430 AGGAATCCCTGTTCTTCAGAGGATGTTGCATATTAAGATTGAGCAAAACATATAAAG 489
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      490 AGAATACAATGACTGTATTCCTCTCGGGCGGTATAGATGCTGGGTACAAACATCAAAA 549
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      550 TTATTCCTCTGTATATTTTGTGGCAAGACTCGTTTCTCTAAGCTGTGCTGATAT 609
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      610 TCTGCATATATAGTTCAGTCGAGCATGATCTAAGTGTGATGGGGTTGATGGA 669
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      670 GTTAGTCAAGCCCAAGCAACTTCTCTCCCTGATATGATAGACTGCGCATTTGGAGCA 729
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      730 AAATAGAGATCACTGCTATCTTGTATTACAGCTTGTGCTGGGCCAAAATTTCTCAAT 789
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      790 GGCATTGATTCAGGCAAGATTCATTCAGAAATATAGAGGACATTGTTATGAGCAAA 849
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      60 ----- 60
DB      850 ATACCTTCGCAATCACTGATGCTGTCGCGAGAAATACCAAAATGAGTGGAGAG 909
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      61 ----- 61
DB      910 AGAGTAGACACAGTCTCTCTTGAATGACAGCTTGCCTTCATTAAGCTAAGATCAG 969
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      64 Ser ----- 64
DB      970 AGTAGGAAAGATGTCTCGATTGAGTGGCCACAAATCCACATCCAGCTCGAGCTCGG 1029
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      64 ----- 64
DB      1030 CAGCAAGTAGCAAGTCCATATCTGACAGAGAAAGTTGGGTCAACGAAAATCTCTTAT 1089

```

```

QY      64 ----- 64
DB      1090 ACTTATTTTCATTAATGACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1149
QY      65 ThrGlyMetValIlePheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGluVal 84
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1150 ACAGAAATATGATTTTCACTACAGATACATATCAACAAATTCAGAGAGCTGAGTA 1209
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      85 ArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPheLysGlyLeuGln 104
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1210 ACTGAAAGCTTTTCTTGTCCATCTGCTAGTAAATGTGCGAGTTTCAGGGCTGAGA 1269
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      105 PheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGluLys 124
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1270 TATCTCTGCTCCATCAACCCAGCATCTCTCAATTCGAGTTTGGGTACGACAAATTT 1329
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      125 GlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGluLysSerAsp 144
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1330 CAGCGGTAATATGCTCCCTCAAGACTGAGACATGATATCCAGTTAATGAGATGAC 1389
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      145 ---AspArgLysPheGluProPheSerLeuCysSerLysPro---ArgLysArgArgGln 162
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1390 GTTGACCCCAAGCAACCAACTTCTTTTCTTCCAAAAATTCAGAGCGAGGCA 1449
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      163 ArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAspSerPro 182
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1450 AAGACTCAGTACGAGCTCAAGGCA -----GGGCT 1482
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      183 SerLeuThrAsnGly-----ThrGluAsnGlyLysThrLeuLeuAsn 196
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1483 CATCTTGATAGTTAGTTGCGAGGTGCTAGATAGACATGATGATGATGATCTGTGATGA 1542
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      197 AspGlyAsnArg-----GlyLeuGlyTyrProGluAlaThrGlyLeuAlaGlyGln 213
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1543 GAGAGAGCCGATACCACTGGAAGAGCATTAC-----GAAAGAAATTTGGGGCT 1590
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      214 PheGluMetThrSerAsnIlePro-----ProAlaIleAlaHisSerSerLeu 229
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1591 GCTGAGTCTGCGCAAGAGTCTCTCTGCGACAGATCTCTGCGAGAGCTGCAATCATGTGG 1650
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      230 AspAlaGlyAlaLysValIleLeuThrSerGluAlaValAlaProAlaThrLysThrArg 249
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1651 GATCCAGATTAATGCTACGATACGATGGAATGCAATGTTGCAAAAACGAG 1710
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      250 LysLeuSerAlaGlnArgSerGlnAlaArgSerHisLeuLeuGlnLysArgGlnPhe 269
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1711 AAAATATCTATAGAACGCTGCGACTTGAGGAACCGAAGCCTCTCAGAGAGACAGTTC 1770
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      270 TyrHisSerHisArgValIleGlnProMetAlaLeuGlnGlnValMetSerAspArgAspSer 289
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1771 TTCACCTCTCATGAGCTCAGCCCATGGCTCTAGAACAGATCTTTCGACCGGATAGT 1830
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      290 GluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPhe 309
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1831 GAAAGATGAAGTTGATGATGATGTGGCAGATTGTAAGATAGAAGATGCTCATGATTTTC 1890
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      310 ValAspValAsnLysAspGlnLysGlnPheMetHisLeuThrAsnSerPheValArgLys 329
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1891 GTTGATGTGACTAAAGATGAGAACAGATGACACATGCACTGTGTTGTGAGAG 1950
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      330 GlnArgValIleAlaAspGlyHisIleSerThrLysGlnAlaPheSerArgPheArg 349
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      1951 CAGCGAGATTTAGCAGATGTGCTACATTCATGCGCATGCGAGGCACTTCCAAAGATTGCA 2010
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      350 GlnLysGluLeuHisArgTyrSerSerLeuPheThrCysTyrArgLeuPheLeuLys 369
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      2011 GGAACCATCATGTTGGAACACCGCATTTGATTTGCTGAGAGATGTTATAGTGTA 2070
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      370 LeuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAspCysAsnThrIleLeuGln 389
DB      ||||| :|:|:| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      2071 CTGTGAAACCAAGGCTTCTTGATGCTCCGCAACATGAACAACTTCAATCTTCTGAA 2130

```


XX EP67278-A2.
 XX 29-DEC-1999.
 XX 28-JUN-1999; 99EP-0305077.
 XX 26-JUN-1998; 98JP-0180065.
 XX 24-JUN-1999; 99JP-0179043.
 XX (MITA) MITSUI CHEM INC.
 XX Yoshida N, Kato Y, Takahashi S, Yanai Y, Hiratsuka J, Miwa T;
 XX WPI: 2000-064612/06.
 DR P-PSDB: AAY53932.
 XX Novel DNA used to produce transgenic plants with altered floral
 XX regulation which can have increased crop yields -
 XX
 PS Example 1: Page 26-34; 53pp; English.

CC The present sequence encodes a protein having a flowering regulating
 CC activity, which is designated Mpc1. The cDNA sequence is given
 CC in AA236947. A rice Mpc1 is also disclosed in the specification. The
 CC rice and Arabidopsis cDNAs show significant homology with each other.
 CC A naturally occurring mutation of the Mpc1 gene eliminates normal
 CC flowering regulating ability of plants, and leads to flowering
 CC immediately after germination (super early flowering mutation). The
 CC Mpc1 polynucleotide sequence can be used to produce plants with altered
 CC flowering times in comparison with wild type plants, by enhancing
 CC or inhibiting the expression of the flowering regulating gene. Antisense
 CC polynucleotides can be used to reproduce the effects of the mutated Mpc1
 CC gene. This alteration can be used to increase the yield of food crops.
 CC flowering inhibition of vegetables increases their productivity.

XX Sequence 5580 BP; 1496 A; 999 C; 1075 G; 2010 T; 0 other;

Alignment Scores:

Pred. No.: 2.4e-32 Length: 5580
 Score: 407.50 Matches: 187
 Percent Similarity: 24.79% Conservative: 47
 Best Local Similarity: 19.81% Mismatches: 87
 Query Match: 17.14% Indels: 626
 DB: 21 Gaps: 18

US-09-890-220-2 (1-445) x AA236948 (1-5580)

QY 65 ThrglyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrcGluVal 84
 DB 2722 ACAGGAATGTAGTTTCACTACAGATCTATACACAAATTCGACAAAGCTGAAG- 2780
 QY 84 -----
 DB 2781 TAACCTAGTATTATTAACTGTTTCACTACCACTGTCTATATTTCATCCGTTACCC 2840
 QY 85 -----
 DB 2841 TAACCTGTACGTATATGTTTGTCTATGCTCTTCGACGTAACCTGAAGACTTTCTTGTTCCA 2900
 QY 92 PheCysSerMetLeuCysGlySerPheLys-----ArgGluAspCysSerCysPro 91
 DB 2901 TTTCGCTTAGTAATAATGTGCAGTTTCAAGGTGACCTTCATTCCATTCCATTCATCC 2960
 QY 101 -----
 DB 2961 TCTTAGTCAAGATACAGCTGTAGTACTGTTTGTAGTATGCAATCTTTCTTTT 3020
 QY 102 -----
 DB 3021 CTCCTCAATCATGTTGTAGGCGCTGAGATATCATCTTGCCATCAACCCAGATCTCCTCAAT 3080
 QY 116 PheGlu-PheLysIeu----- 120

DB 3081 TTTCAGTTTGGGTTGTAGCTTTAAATTCAGTTAACCTGTTTGCATTTTTTTTANT 3140
 QY 121 -----PheGluIuTyrGlnThrValAs 128
 DB 3141 TTGTGGGTCACCAATATCTGCTTACTTGTAGTAACTGACAAATTCAGGCGTAAA 3200
 QY 128 nValSerValLys----- 132
 DB 3201 TGTCTCCCTCAAGACTGACAGCAATGATTCAGAGGTTAGAACATCTGTTGTCGATTT 3260
 QY 132 ----- 132
 DB 3261 ATGTTCAATAGTTTCTGCTGTAATCTATAGCTTAACAAATTCATTTTCATTTA 3320
 QY 133 -----LeuAsnSerPheIlePheGluGluGlySer----- 143
 DB 3321 AACATAATCCCTCATGGGTTGTGACTTTGTGTGTTAATAAGGAACTGCAATCTT 3380
 QY 144 -----AspAspAsp 147
 DB 3381 TAGTTGCTATTGTCACACTATGATCTTGTCTATGCTTAAATAGCTGATGAGAAATTA 3440
 QY 147 sPheGlu----- 149
 DB 3441 ACT-CAAAATGACATCGCTGTTCTGTTACTTTTGTGGCATGACCGCAAAAGCTCG 3499
 QY 149 ----- 149
 DB 3500 ACTGTAGATAAAGTCTCGATTAATATAGAGTCAATCTAATTAAGATGTTGTTTC 3559
 QY 149 ----- 149
 DB 3560 TACAAATATTTCTATGCTTTGTAGTCTTCTTATTAATGATTAATCTTACACACATTT 3619
 QY 149 ----- 149
 DB 3620 TCTAATGTTAATGTTTCACTTAAATAATGCTGCTCAGGTTAATAGAGTACGTTGACCA 3679
 QY 150 -----ProPhe----- 151
 DB 3680 AGCAGCAAACTTCTTTTTCGTAAATATCTGCGCTATATGTTGCTTTTATATCTT 3739
 QY 152 -----SerLeuCysSerLysProArgLysArgAr 161
 DB 3740 TCCACACTCTGTGTGACACCAATAAATTTCTCAATATGTGACAGTTCCAAAATTTGAG 3799
 QY 161 gGlnArgGlyLysArgAsnAsnThrArgArgLeuLysVal----- 174
 DB 3800 ACGGAGGAGGCAAAAGAGTCAAGGTACGAGGTCAGAGGCAAGGCGCTCATCTGATTAGG 3859
 QY 175 -CysPheLys-----ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGly 192
 DB 3860 TTTCGAGGTGCTAGATAAGACTGAGGTATGTGTGACTGAAATGACAGATTAATGGAT 3919
 QY 192 ethrLeuLys----- 195
 DB 3920 TTGTAGTATTTGGCTTTTGTGATGACAGAGCTGCTTATGTTATTTTACAGATAT 3979
 QY 196 -----AsnAspGlyLysAsnArgGlyLeuGlyTyrPro 205
 DB 3980 TTACTTGTATATGTCATATTTTGCATGCAACAACTGATGATCTTGACAGAGCTTTTAA 4039
 QY 205 oGluAlaIthrGluLeuAlaGlyGlnPheGluMet--ThrSerAsnIleProProAlaIle 224
 DB 4040 TTCTTACTTGTAGATGCTCATCTCTGTAGAACTGAGAAAGCGAATCCACCTGGA-- 4096
 QY 225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValPro 244
 DB 4097 AAGCATTCAGAAAGATTTGGGGGTCTGAG-----TCGTCGCAAGAGTCTCTCT 4147
 QY 245 AlaThr----- 246

Db 4148 GGCACGAGTCCTGCAGACGTCGATCATGATGGGAGATCCAGATATATGTCAGTCGATAGCT 4207
 QY 247 -----LysThrArgLysLeuSerAlaGluArgSerLysAla 258
 Db 4208 GGAAGTACAAATGTTCAGATTTGCAGAAAACGAGAAATATCTATAGAACGCTGCAGACTTG 4267
 QY 259 ArgSerHisLeu----- 262
 Db 4268 AGCAAGTATGTTGACTTCCTTTGCTGCTCATCCCTCTTCTCAATTTATATTAACTA 4327
 QY 262 ----- 262
 Db 4328 CATATGTCATGATGAGAAAAATGTGCTAGTTTATTAACAAGTACCTGTTAAATCCC 4387
 QY 263 -----LeuLeuGlnLys 266
 Db 4388 AATGATGTGAGTGAAGTTTTCAAAATTTTTCCTCCACGCGAAGCCTCCTTCAGAG 4447
 QY 267 ArgGlnPheTyrHisSerHisArgVal----- 275
 Db 4448 AGCAGTCTCTCCACTCATCTCATGACCTCAGCTGAGATCTTTTCTTAGCTCTCTGCTT 4507
 QY 276 -----GlnProMetAlaLeuGlu 281
 Db 4508 TGAAGATTCATGATGATTTTTCATTTGCTATGCTACTATGCGACGCCCATGCGCTAGAA 4567
 QY 282 GlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGlu 301
 Db 4568 CAAGTACTTTCGACCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 4627
 QY 302 AspArg----- 303
 Db 4628 GATAGAGAGGCTATGTTTGTGATTTATATTTTTCACCGCATCACTAGTTGGGTGAAATA 4687
 QY 303 ----- 303
 Db 4688 AGCTCAGTAGTTGGGTAGATATATGTTTCATGTCGAAAGGAGAAAGAAATATTGAAGCTGG 4747
 QY 303 ----- 303
 Db 4748 GCATGGCGCAACGTTAGAGCAATATTTGTAATGTTGACAGATCAATAGAAAAATATCTGA 4807
 QY 303 ----- 303
 Db 4808 GCAAGCCTCAGCGTTGATATGAGAACAGTAGAACAGATCATAGTCTTATATACACT 4867
 QY 303 ----- 303
 Db 4868 CATTAAGACGAGAGTGTCCGTTGTACTCGATCTCAACATAGTTGATTCATACATAG 4927
 QY 304 -----GlnMetLe 306
 Db 4928 TTGTCTCATCTCCATATAGTGAATACGTTATTTCTTATTTCTTCTTCAGATGCT 4987
 QY 306 UnspAspPheValAspValAspValAspGluLysGlnPheMetHisLeuTrpAsnSerPhe 326
 Db 4988 CGATGATTTTCGTTGCTGATGACTAAAGATGAGAACATGATGATGATGATGATGATGATGAT 5047
 QY 326 eValArgLysGlnArgVal----- 332
 Db 5048 TGTAGAGAAAGCAGCGGTA-TGCTTATCTCTTTTCATGATGATGATGATGATGATGATGAT 5106
 QY 332 ----- 332
 Db 5107 GTATTAACATTTAGAGTGGCGCATGTAAGGTTTGATGATTAATTCCTGCTGGCTTCTCT 5166
 QY 332 ----- 332
 Db 5167 GGTAAAAAATAAACTGAACAATGATTAACATGACATGCTCTCTGACTACT 5226
 QY 333 -----IleAlaAspGlnHisIle 338
 Db 5227 ATTAAGCATTAACCTTGACAGTGTGTTTGACCCCTTTCAGAGATATTAAGCATGATGTCACAT 5286

QY 338 eSerTrpAlaCysGlnAlaPheSerArgPheTyrGlnLysGlnLeuHisArgTyrSer 358
 Db 5287 TCATAGGACATGACGAGGATTTCTCAATATTCACAGGACCATCATCATGTTGCAACACGCA 5346
 QY 358 IleuPheTrp----- 361
 Db 5347 CTGTATTTGTAATTAACACTCTCATTTCTTCATTTGTTTTCAGAGTATGCGAAGA 5406
 QY 362 -----CysTrpArgLeuPheLeuIle 368
 Db 5407 AAGCGTTTGTGATTAAGTAGACTTTTGTGTGATGCTGCTGAGAGTGTATATGCT 5466
 QY 368 eLysLeuTrpAsnHisGlnLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLe 388
 Db 5467 GAACCTGTGAACACGCTCTTGTATGCCGCAACCATGACAACTGTAATATACCTTCT 5526
 QY 388 uGlu 389
 Db 5527 CGAA 5530
 Db 5527 CGAA 5530
 RESULT 13
 AAA47757
 ID AAA47757 standard; cDNA: 325 BP.
 AC AAA47757;
 DT 16-NOV-2000 (first entry)
 DE A1163743 Expressed sequence tag (modified) cDNA.
 XX Vernalization gene; VRN2; plant characteristic; flowering time; poplar;
 KW leaf size; leaf shape; shade avoidance response; reproduction; EST;
 KW breeding; pollination; cultivation; expressed sequence tag; ss.
 OS Populus tremula.
 PN WO20044918-A1.
 PD 03-AUG-2000.
 PF 28-JAN-2000; 2000WO-GB00248.
 PR 28-JAN-1999; 99GB-0001927.
 PA (PLAN-) PLANT BIOSCIENCE LTD.
 PI Dean C. Gendall A;
 DR WPI; 2000-499333/44.
 DR P-PSDB; AAB00064.
 PT Isolated vernalization gene VRN2 is used to produce transgenic plants
 PT with altered vernalization response, flowering time, leaf size and/or
 PT shape or shade avoidance response for maximized reproductive success
 PS Disclosure; Page 76; 105pp; English.
 CC Isolated nucleic acid sequences obtained from the VRN2 locus of a
 CC plant, encode polypeptides which are capable of affecting one or more
 CC vernalization responses such as, flowering time, leaf size and/or
 CC shape or the shade avoidance response of a plant into which the
 CC nucleic acid is introduced. Introducing such sequences into plants
 CC to alter these characteristics maximizes the reproductive success of
 CC the plant. This cDNA represents expressed sequence tag clone A1163743
 CC from poplar trees (Populus tremula x Populus tremuloides) which has
 CC homology in the C-terminal region of the Arabidopsis thaliana VRN2 gene.
 SO Sequence 325 BP; 102 A; 60 C; 76 G; 87 T; 0 other;
 Alignment Scores: 4,84e-30 Length: 325
 Pred. No.: 368.50 Matches: 76
 Score:

PR	21-JUL-1999;	99US-0144814
PR	21-JUL-1999;	99US-0145096

```

PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145313.
PR 27-JUL-1999; 99US-0145318.
PR 27-JUL-1999; 99US-0145319.
PR 28-JUL-1999; 99US-0145351.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147360.
PR 06-AUG-1999; 99US-0147303.
PR 09-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148319.
PR 13-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148341.
PR 16-AUG-1999; 99US-0148364.
PR 17-AUG-1999; 99US-0148375.
PR 18-AUG-1999; 99US-0148426.
PR 20-AUG-1999; 99US-0148722.
PR 20-AUG-1999; 99US-0148723.
PR 23-AUG-1999; 99US-0149329.
PR 23-AUG-1999; 99US-0149302.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.

```

```

PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:
Pred. No.: 1,29e-26
Score: 340.50
Percent Similarity: 80.00%
Best Local Similarity: 78.82%
Query Match: 14.32%
DB: 21 Gaps: 2

US-09-890-220-2 (1-445) x AAC44147 (1-640)
OY 362 CysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIle 381
Db 305 TCGTGGCGG-----TGG-----TGC 319
OY 382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
Db 320 AACCACTGCATACCATCTCGAGAAATTGCCGTAAATAGCTCAGACACCCACCAAC 379
OY 402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAsp 421
Db 380 AACCAACACAGCTGATCGTCCAGACTCAACACCAACCAACAAATTAATCTGAT 439
OY 422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
Db 440 CATCCCAATGACATTAACAAACAAAGAACAAATGTTGACACAAAGACAAATTAACGACAGAA 499
OY 442 -LysValIleLys 445
Db 500 CAAAGTAATTAA 512

RESULT 15
AAC51582
ID AAC51582 standard; DNA; 319 BP.
XX
AC AAC51582;
XX
DI 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 69042.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.

```


PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147034.
PR 04-AUG-1999; 99US-0147208.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158360.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 13-OCT-1999; 99US-0158293.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159584.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Alignment Scores:

Prd. No.: 1,259-25
Score: 327.00
Percent Similarity: 96.838
Best Local Similarity: 96.838
Query Match: 13,758
Db: 21
Length: 319
Matches: 61
Conservative: 0
Mismatches: 2
Indels: 0
Gaps: 0

US-09-890-220-2 (1-445) x AAC51582 (1-319)

QY 383 AsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspPrhThrThrAsnAsn 402
Db 3 AACCTGCAATACCATCTCGAGATTCGCGTAATAGCTCAGACACCACCCACCAAC 62
QY 403 AsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHis 422
Db 63 AACACACAGTGTGATGTCCTCCAGTCACCAACACCAACATTAACATTGTGATCAT 122
QY 423 ProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSerArgAspLys 442
Db 123 CCCAATGACATTAACACACAGAACATGTTGACAAACACATTAACAGCAGACACAA 182
QY 443 ValIleLys 445
Db 183 GTAATTAA 191

Search completed: June 19, 2003, 23:51:41
Job time : 377 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:39:38 ; Search time 70 Seconds
(without alignments)
1949.588 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONRAKSSPEEYISTDE.....INKNKNVDKNNSRDYIK 445

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 segs, 15338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_model -DEV=xlh
-c=/cgn2_1/USPTO_spool/US09890220/runat_13062003_144852_16502/app_query.fasta.1.583
-DB=Issued_Patents_NA -QFMT=fastap -SUFFIX=ini -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09890220.ecgn.1.140.etruncat_13062003_144852_16502 -NCPD=6 -ICPD=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Issued Patents_NA:*
2: /cgn2_6/ptodata/1/ina/5A.COMB.seg:*
3: /cgn2_6/ptodata/1/ina/5B.COMB.seg:*
4: /cgn2_6/ptodata/1/ina/6A.COMB.seg:*
5: /cgn2_6/ptodata/1/ina/6B.COMB.seg:*
6: /cgn2_6/ptodata/1/ina/Backfiles1.seg:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	120.5	5.1	198	PCT-US95-10668-3	Sequence 3, Appl1
2	120.5	5.1	198	PCT-US95-10668-4	Sequence 4, Appl1
3	114	4.8	198	PCT-US95-10668-1	Sequence 1, Appl1
4	114	4.8	198	PCT-US95-10668-2	Sequence 2, Appl1
5	109	4.6	2237	US-08-914-999-7	Sequence 7, Appl1
6	107	4.5	1086	US-08-415-751-27	Sequence 27, Appl1
7	107	4.5	1086	US-08-415-751-28	Sequence 28, Appl1
8	107	4.5	5163	US-08-700-651-1	Sequence 1, Appl1
9	107	4.5	5163	US-08-928-361B-4	Sequence 4, Appl1
10	107	4.5	5318	US-08-700-651-2	Sequence 2, Appl1
11	107	4.5	5318	US-08-928-361B-3	Sequence 3, Appl1
12	107	4.5	5511	US-08-928-361B-2	Sequence 2, Appl1

13	107	4.5	7334	3	US-08-928-361B-1	Sequence 1, Appl1
14	105.5	4.4	1430	1 <td>US-08-276-452A-25</td> <td>Sequence 25, Appl1</td>	US-08-276-452A-25	Sequence 25, Appl1
15	105.5	4.4	1430	2 <td>US-08-798-744-25</td> <td>Sequence 25, Appl1</td>	US-08-798-744-25	Sequence 25, Appl1
16	104.5	4.4	1680	2 <td>US-08-700-152A-3</td> <td>Sequence 3, Appl1</td>	US-08-700-152A-3	Sequence 3, Appl1
17	103	4.3	4223	4 <td>US-09-541-782-5</td> <td>Sequence 5, Appl1</td>	US-09-541-782-5	Sequence 5, Appl1
18	103	4.3	4223	4 <td>US-09-723-820-5</td> <td>Sequence 5, Appl1</td>	US-09-723-820-5	Sequence 5, Appl1
19	98.5	4.1	4757	4 <td>US-09-572-191-1</td> <td>Sequence 1, Appl1</td>	US-09-572-191-1	Sequence 1, Appl1
20	98.5	4.1	4757	4 <td>US-09-723-262-1</td> <td>Sequence 1, Appl1</td>	US-09-723-262-1	Sequence 1, Appl1
21	98.5	4.1	4757	4 <td>US-09-723-219-1</td> <td>Sequence 1, Appl1</td>	US-09-723-219-1	Sequence 1, Appl1
22	98.5	4.1	7425	4 <td>US-09-453-702B-212</td> <td>Sequence 212, App</td>	US-09-453-702B-212	Sequence 212, App
23	98	4.1	1086	1 <td>US-08-415-751-47</td> <td>Sequence 47, App</td>	US-08-415-751-47	Sequence 47, App
24	97	4.1	1690	1 <td>US-08-276-452A-24</td> <td>Sequence 24, Appl1</td>	US-08-276-452A-24	Sequence 24, Appl1
25	97	4.1	1690	2 <td>US-08-798-744-24</td> <td>Sequence 24, Appl1</td>	US-08-798-744-24	Sequence 24, Appl1
26	96.5	4.1	2093	1 <td>US-08-287-001A-1</td> <td>Sequence 1, Appl1</td>	US-08-287-001A-1	Sequence 1, Appl1
27	96.5	4.1	2093	1 <td>PCT-US95-09941-1</td> <td>Sequence 1, Appl1</td>	PCT-US95-09941-1	Sequence 1, Appl1
28	96	4.0	1956	4 <td>US-08-559-896B-1</td> <td>Sequence 1, Appl1</td>	US-08-559-896B-1	Sequence 1, Appl1
29	94	4.0	8920	2 <td>US-08-446-855A-1</td> <td>Sequence 1, Appl1</td>	US-08-446-855A-1	Sequence 1, Appl1
30	94	4.0	8920	4 <td>US-09-150-741-1</td> <td>Sequence 1, Appl1</td>	US-09-150-741-1	Sequence 1, Appl1
31	93	3.9	700	4 <td>US-09-236-097-11</td> <td>Sequence 11, Appl1</td>	US-09-236-097-11	Sequence 11, Appl1
32	93	3.9	5720	4 <td>US-09-442-100-1</td> <td>Sequence 1, Appl1</td>	US-09-442-100-1	Sequence 1, Appl1
33	91.5	3.8	2136	4 <td>US-09-177-248-1</td> <td>Sequence 1, Appl1</td>	US-09-177-248-1	Sequence 1, Appl1
34	91.5	3.8	2315	4 <td>US-09-061-769A-1</td> <td>Sequence 1, Appl1</td>	US-09-061-769A-1	Sequence 1, Appl1
35	91.5	3.8	2315	4 <td>US-09-061-769A-4</td> <td>Sequence 4, Appl1</td>	US-09-061-769A-4	Sequence 4, Appl1
36	91.5	3.8	2327	4 <td>US-09-061-769A-3</td> <td>Sequence 3, Appl1</td>	US-09-061-769A-3	Sequence 3, Appl1
37	91.5	3.8	5361	4 <td>US-08-973-462-2</td> <td>Sequence 2, Appl1</td>	US-08-973-462-2	Sequence 2, Appl1
38	91.5	3.8	6152	4 <td>US-08-973-462-1</td> <td>Sequence 1, Appl1</td>	US-08-973-462-1	Sequence 1, Appl1
39	91	3.8	1631	3 <td>US-09-118-319-1</td> <td>Sequence 1, Appl1</td>	US-09-118-319-1	Sequence 1, Appl1
40	91	3.8	2288	2 <td>US-08-568-459A-5</td> <td>Sequence 5, Appl1</td>	US-08-568-459A-5	Sequence 5, Appl1
41	91	3.8	2288	2 <td>US-08-487-826B-5</td> <td>Sequence 5, Appl1</td>	US-08-487-826B-5	Sequence 5, Appl1
42	91	3.8	2288	4 <td>US-09-210-288-5</td> <td>Sequence 5, Appl1</td>	US-09-210-288-5	Sequence 5, Appl1
43	91	3.8	3800	4 <td>US-08-965-762-13</td> <td>Sequence 13, Appl1</td>	US-08-965-762-13	Sequence 13, Appl1
44	91	3.8	3800	4 <td>US-09-911-927-13</td> <td>Sequence 13, Appl1</td>	US-09-911-927-13	Sequence 13, Appl1
45	91	3.8	3800	4 <td>US-09-911-927-15</td> <td>Sequence 15, Appl1</td>	US-09-911-927-15	Sequence 15, Appl1

ALIGNMENTS

RESULT 1
PCT-US95-10668-3
Sequence 3, Application PC/TUS9510668
GENERAL INFORMATION:
APPLICANT: James Eberwine
TITLE OF INVENTION: A Method of Sequencing Proteins by
TITLE OF INVENTION: Epitope Ordering and Protein
TITLE OF INVENTION: Restriction Mapping
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jane Massey Licata, Esq.
STREET: 210 Lake Drive East, Suite 201
CITY: Cherry Hill
STATE: NJ
COUNTRY: USA
ZIP: 08002
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
COMPUTER: IBM 486
OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
SOFTWARE: WORDPERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10668
FILING DATE: Herewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/294,133
FILING DATE: August 22, 1994
ATTORNEY/AGENT INFORMATION:
NAME: Jane Massey Licata
REGISTRATION NUMBER: 32,257
REFERENCE/DOCKET NUMBER: PENN-0137
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 779-2400
TELEFAX: (609) 779-8488
INFORMATION FOR SEQ ID NO: 3:

Fri Jun 20 09:03:40 2003

us-09-890-220-2.rml

Page 2

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 198
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
; PCT-US95-10668-3

Alignment Scores:
Pred. No.: 4.91e-06 Length: 198
Score: 120.50 Matches: 24
Percent Similarity: 65.08% Conservative: 17
Best Local Similarity: 38.10% Mismatches: 21
Query Match: 5.07% Indels: 1
DB: Gaps: 1

US-09-890-220-2 (1-445) x PCT-US95-10668-3 (1-198)
QY 382 AsnaAncysAsnThrIleuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
DB 5 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 64
QY 402 AsnaAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAsp 421
DB 65 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 121
QY 422 HisProAsnAspIleAsnAsnValAspAsnValAspAsnValAspAsnSerArgAsp 441
DB 122 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 181
QY 442 LysValIle 444
DB 182 AAGCTGATG 190

RESULT 2
PCT-US95-10668-4
; Sequence 4, Application PC/TUS9510668
; GENERAL INFORMATION:
; APPLICANT: James Eberwine
; TITLE OF INVENTION: A Method of Sequencing Proteins by
; TITLE OF INVENTION: Epitope Ordering and Protein
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10668
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/294,133
; FILING DATE: August 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: PENN-0137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198
; TYPE: Nucleic Acid
```

```

; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
; PCT-US95-10668-4

Alignment Scores:
Pred. No.: 4.91e-06 Length: 198
Score: 120.50 Matches: 24
Percent Similarity: 65.08% Conservative: 17
Best Local Similarity: 38.10% Mismatches: 21
Query Match: 5.07% Indels: 1
DB: Gaps: 1

US-09-890-220-2 (1-445) x PCT-US95-10668-4 (1-198)
QY 382 AsnaAncysAsnThrIleuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
DB 5 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 64
QY 402 AsnaAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleValAsp 421
DB 65 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 121
QY 422 HisProAsnAspIleAsnAsnValAspAsnValAspAsnValAspAsnSerArgAsp 441
DB 122 AACCAACACACACACACACACACACACACACACACACACACACACACACACAC 181
QY 442 LysValIle 444
DB 182 AAGCTGATG 190

RESULT 3
PCT-US95-10668-1
; Sequence 1, Application PC/TUS9510668
; GENERAL INFORMATION:
; APPLICANT: James Eberwine
; TITLE OF INVENTION: A Method of Sequencing Proteins by
; TITLE OF INVENTION: Epitope Ordering and Protein
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jane Massey Licata, Esq.
; STREET: 210 Lake Drive East, Suite 201
; CITY: Cherry Hill
; STATE: NJ
; COUNTRY: USA
; ZIP: 08002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE
; COMPUTER: IBM 486
; OPERATING SYSTEM: WINDOWS FOR WORKGROUPS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10668
; FILING DATE: Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/294,133
; FILING DATE: August 22, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Jane Massey Licata
; REGISTRATION NUMBER: 32,257
; REFERENCE/DOCKET NUMBER: PENN-0137
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 779-2400
; TELEFAX: (609) 779-8488
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 198
; TYPE: Nucleic Acid
; STRANDEDNESS: Single
; TOPOLOGY: Linear
; ANTI-SENSE: No
```



```

OY 261 HistLeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnPrometAlaLeu 280
DB 620 CATTAAACCTCCAAAAAGATTGAAATTTTCAACATGCG-GTT-----TTA 666
OY 281 GlnGlnValMetSerAspArgAspSerGluAspGluVal-----AspAspArgValAla 298
DB 667 GAGTTTGTGTA-----GATGACACATCATCATGATTTTAAATTTGTGAGCAGCAACATGTGA 720
OY 299 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAspGlnLysAspGlnLys 317
DB 721 GAAGGACAAATATGAAAGTATATATATATATATATATATATATATATATATATATATATAT 780
OY 318 -----GlnPheMetHisLeuThrPasnSerPheValArgLysGlnArgValIleAla 334
DB 781 ACACACACATCATCTCTCTCATTTCCATATGAAACATCAATCAATCAATTTATGATTATA 840
OY 335 Asp-----GlnHisIleSerThrPalaCysGluAlaPheSerArgPheThrGlu 350
DB 841 GATATTCAGGTGTGGT-----GATCCTATATACAGC 873
OY 351 LysGlnLeuHisArgTyrSerSerLeuPheThrPyrGluPheLeuIleLysLeu 370
DB 874 CCACAAATTCATACCTATG-----TGCTGTGGT-TTGCTATGTGGTAT-TTG 921
OY 371 TrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlu 389
DB 922 GGTCAAAAAGGTTTGAAGTTTATGACTCTCAATGATATGCAATTTGCCAATAT 981
OY 390 ---AsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAspArg 408
DB 982 TTAATTTTCAATCAATTAATGCAAAATCTGAAAAAGGATTTGTGATGCTGCTCAACAA 1041
OY 409 ProSer-----AspSerAsnThrAsnAsnAsnAsnIle 419
DB 1042 CCAGATTTAAATTTCCCTGATCATCTGAAAGATATATATATATATATATATATATATAT 1098
OY 420 ValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSer 439
DB 1099 ---AAATATATATATATATATATATATATATATATATATATATATATATATATATAC 1155
OY 440 ArgAspLysValIleLys 445
DB 1156 AATAGTAGTATTTCAAAA 1173

RESULT 6
US-08-415-751-27
Sequence 27, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PHILLIPS, MOORE, LEMPJO & FINLEY
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 KB storage
COMPUTER: PC
OPERATING SYSTEM: DOS
SOFTWARE: WordPerfect 5.1

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,751
FILING DATE: 03-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/071,880
FILING DATE: June 1, 1993
APPLICATION NUMBER: 07/891,301
FILING DATE: May 29, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Hana Dolezalova
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.19-2 (HRD)
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 324-1677
TELEFAX: (415) 324-1678
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 1086 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Cryptosporidium parvum

US-08-415-751-27

Alignment Scores:
Pred. No.: 0.00383 Length: 1086
Score: 107.00 Matches: 23
Percent Similarity: 58.33% Conservative: 12
Best Local Similarity: 38.33% Mismatches: 25
Query Match: 4.50% Indels: 0
Gaps: 0

US-09-890-220-2 (1-445) x US-08-415-751-27 (1-1086)

OY 382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
DB 345 AACACACACACAACTCTATCCATTCATTCACACACACACAAACCAACAAACAA 404
OY 402 AsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAsp 421
DB 405 AACACACACACAAAGGTCAGCTAGCCACCAATAGCCACACACACAAACATTAAGGC 464
OY 422 HisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerAsp 441
DB 465 AATAGTTTCAACACACACAAACAAACACACACACACACACACACACAGTCCACAGC 524

RESULT 7
US-08-415-751-28/C
Sequence 28, Application US/08415751
Patent No. 5643772
GENERAL INFORMATION:
APPLICANT: PETERSEN, CAROLYN
APPLICANT: LEECH, JAMES
APPLICANT: NELSON, RICHARD, C.
APPLICANT: GUT, JIRI
TITLE OF INVENTION: POLYPEPTIDES BINDING ANTI-
TITLE OF INVENTION: CRYPTOSPORIDIUM ANTIBODIES, DNA
TITLE OF INVENTION: AND RNA ENCODING THEM, HYBRID
TITLE OF INVENTION: VECTOR AND TRANSFORMED HOST AND
TITLE OF INVENTION: METHODS FOR IMMUNOTHERAPY AND
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PHILLIPS, MOORE, LEMPJO & FINLEY
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: California
COUNTRY: United States of America
ZIP: 94306-1840
COMPUTER READABLE FORM:

```

MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Kb storage
 COMPUTER: PC
 OPERATING SYSTEM: DOS
 SOFTWARE: Wordperfect 5.1
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/415,751
 FILING DATE: 03-APR-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/071,880
 FILING DATE: June 1, 1993
 APPLICATION NUMBER: 07/891,301
 FILING DATE: May 29, 1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Hana Dolezalova
 REGISTRATION NUMBER: 30,518
 REFERENCE/DOCKET NUMBER: 480.19-2 (HHD)
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (415) 324-1677
 TELEFAX: (415) 324-1678
 INFORMATION FOR SEQ ID NO: 28:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1086 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 ORIGINAL SOURCE:
 ORGANISM: *Cryptosporidium parvum*
 US-08-415-751-28


```

TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verny, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480.76-1.(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5511 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-928-361B-2

Alignment Scores:
Pred. No.:      Length:     5511
Score:          Matches:    23
Percent Similarity: 58.33%   Conservative: 12
Best Local Similarity: 38.33% Mismatches:    25
Query Match:      Gaps:       Indels:        0
DB:                3           0

US-09-890-220-2 (1-445) x US-08-928-361B-2 (1-5511)
QY      382 ASnAsnCysGaspnrhlleuGuIsAcNsGsArGaSnSeSerApsrthrtrhnAsn 401
         ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2163 AACACACCACAATTCCTTAATCCCATCACTTACACAAACGCCAAAAACCAACAGC 2222
QY      402 AsnAsnAsnServalAsparGProSerAapSerAasrThraSnAsnAsnAnlleValAsp 421
         ||||||| || |:::||| ||||| ||||| ||||| ||||| ||||| |||||
Db      2223 AACACAACAAGAAGTGCGAGGTAAAGCACCAATAGCCAACACAACAATTTAAAGCC 2282
QY      422 HisPrGaaSplLeAsnsnlysaSaasnValAspAsnllyspAsnaSnbnseratgaSp 441
         ::::: ||| ||||| ||:::||| :||| :||| ||::: |||
Db      2283 AATAgTTACAACAACAACAACAACAACAACAACAACAACAACAAGTCCAAGCAC 2342

RESULT 13
US-08-928-361B-1
Sequence 1, Application US/08928361B
Patent No. 6071518
GENERAL INFORMATION:
APPLICANT: Petersen, Carolyn
TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,
TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS
TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM
TITLE OF INVENTION: SPECIES INFECTIONS
NUMBER OF SEQUENCES: 30
CORRESPONDENCE ADDRESS:
ADDRESSEE: PETERS, VERNY, JONES & BIKSA
STREET: 385 Sherman Avenue, Suite 6

```

```

CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94306-1840

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentln Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/928,361B
FILING DATE: 12-SEP-1997
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/026,062
FILING DATE: 13-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Verry, Hana
REGISTRATION NUMBER: 30,518
REFERENCE/DOCKET NUMBER: 480,76-1(HV)
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-324-1677
TELEFAX: 650-324-1678
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 7334 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-928-361B-1

Alignment Scores:
Pred.: 0.0953 Length: 7334
Score: 107.00 Matches: 23
Percent Similarity: 58.33% Conservative: 12
Best Local Similarity: 38.33% Mismatches: 25
Query Match: 4.50% Indels: 0
DB: 3 Gaps: 0

US-09-890-220-2 (1-445) x US-08-928-361B-1 (1-7334)
QY 382 AsnaNsCysAsnTrIleLeuGlnAsnCysArgAsnSerSeraSPThrThrThraSn 401
||||| ||| ||| ||||| |||
Db 3831 AACCAACCAACAATCTTTAATCCATCATTTACAAACAACAACCAACAAC 3890
QY 402 AasnasnservValaspPProseraspsersantnrtnAasnasnasnil evalasp 421
||||| ||| ||| :::: ||||||||
Db 3891 AACCAACAACAAGTTCACAGTAAAGCCACCATAATGCCCAACAACAACAACATTAAGCC 3950
QY 422 HisProAsnspIlleasnAnsnLysAsnAsnValaspasnLysaspAsnAsnSerArgasp 441
::: |||||:::||| ::||| ::|||
Db 3951 AATAGTTACAAACAACAACAACAACAACAACAACAACAACAACAGTGCCAACGAC 4010

RESULT 14
US-08-276-452A-25
Sequence 25, Application US/08276452A
Patent No. 5646029
GENERAL INFORMATION:
APPLICANT: Chen, Chao-Guang
APPLICANT: Mau, Shiao-Lim
APPLICANT: Du, He
APPLICANT: Gane, Alison M
APPLICANT: Bacic, Antony
APPLICANT: Clarke, Adrienne E
TITLE OF INVENTION: Plant Arabidnagalactan Protein (Agp) Genes
NUMBER OF SEQUENCES: 91
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: Colorado
COUNTRY: United States of America

```

```

? ZIP: 80303
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/276,452A
? FILING DATE: 18-JUL-1994
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Caruthers, Jennie M.
? REGISTRATION NUMBER: 34,464
? REFERENCE/DOCKET NUMBER: 27-91A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (303)499-8080
? TELEFAX: (303)499-8089
? TELETYPE: 49617824
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1430 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 2..1312
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 41..112
? OTHER INFORMATION: /note= "Derived amino acid sequence
? OTHER INFORMATION: corresponding to the peptide sequence by protein
? OTHER INFORMATION: microsequencing"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 41..112
? OTHER INFORMATION: /note= "Derived amino acids 14-24,
? OTHER INFORMATION: 28, 30, 32-37 are identical to sequences obtained
? OTHER INFORMATION: by protein microsequencing"
? NAME/KEY: misc_feature
? LOCATION: 25..31
? OTHER INFORMATION: /note= "Amino acids 25, 27, 29, and
? OTHER INFORMATION: 31 are hydroxylated proline residues; amino acid
? OTHER INFORMATION: 26 can be T instead of A"
?
? US-08-276-452A-25
?
? Alignment Scores:
? Pred. NO.: 0.00927 Length: 1430
? Score: 105.50 Matches: 28
? Percent Similarity: 42.11% Conservative: 12
? Best Local Similarity: 29.47% Mismatches: 32
? Query Match: 4.44% Indels: 23
? DB: 1 Gaps: 2
?
? US-09-890-220-2 (1-445) x US-08-276-452A-25 (1-1430)
?
? QY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlnAsnCys 391
? DB 551 AACATGCTACTCGAGATTTACAAACAACAACAATGGCTACTCGAGATTTAC 610
? QY 392 ArgAsnSerSerAsp-----ThrThrThrAsnAsnAsnSerValAspArgProSer 410
? DB 611 AACACAAACAACAATGGCTACTCGAGATTTACAAACAACAATGGCTACTCGAGATTTAC 610
? QY 397 -----ThrThrThrAsnAsnAsnSerValAspArgProSer 410
? DB 671 AACAAATATGGTTACTCCAGATTTACATGATGACAAACAACGGCTCTCCGAGAGTTAC 730
? QY 411 AspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnIleAsn 430
? DB 731 AACAGCAACAACAACAACAATATTTCTCCGAGATTTACAAACAACAATATTAACAAT 730

```

```

? QY 431 AsnVal-----AspAsnIysAspAsnAsnSerArgAspIysVal 443
? DB 791 AATGTTTCTCCGAGATTTACAAACAACAACAATATTAACAATATGTT 835
?
? RESULT 15
? US-08-798-744-25
? Sequence 25, Application US/08798744
? Patent No. 5830747
? GENERAL INFORMATION:
? APPLICANT: Chen, Chao-Guang
? APPLICANT: Mau, Shiao-Lim
? APPLICANT: Du, He
? APPLICANT: Gane, Alison M
? APPLICANT: Bacic, Antony
? APPLICANT: Clarke, Adrienne E
? TITLE OF INVENTION: Plant Arabidnagalactan Protein (Acp) Genes
? NUMBER OF SEQUENCES: 91
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Greenlee, and Winner, P.C.
? STREET: 5370 Manhattan Circle, Suite 201
? CITY: Boulder
? STATE: Colorado
? COUNTRY: United States of America
? ZIP: 80303
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/798,744
? FILING DATE: 13-FEB-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/276,452
? FILING DATE: 18-JUL-1994
? ATTORNEY/AGENT INFORMATION:
? NAME: Caruthers, Jennie M.
? REGISTRATION NUMBER: 34,464
? REFERENCE/DOCKET NUMBER: 27-91A
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (303)499-8080
? TELEFAX: (303)499-8089
? TELETYPE: 49617824
? INFORMATION FOR SEQ ID NO: 25:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1430 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 2..1312
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 41..112
? OTHER INFORMATION: /note= "Derived amino acid sequence
? OTHER INFORMATION: corresponding to the peptide sequence by protein
? OTHER INFORMATION: microsequencing"
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: 41..112
? OTHER INFORMATION: /note= "Derived amino acids 14-24,
? OTHER INFORMATION: 28, 30, 32-37 are identical to sequences obtained
? OTHER INFORMATION: by protein microsequencing"
? NAME/KEY: misc_feature
? LOCATION: 25..31
? OTHER INFORMATION: /note= "Amino acids 25, 27, 29, and
? OTHER INFORMATION: 31 are hydroxylated proline residues; amino acid
? OTHER INFORMATION: 26 can be T instead of A"

```

US-08-798-744-25

Alignment Scores:

Pred. No.:	0.00927	Length:	1430
Score:	105.50	Matches:	28
Percent Similarity:	42.11%	Conservative:	12
Best Local Similarity:	29.47%	Mismatches:	32
Query Match:	4.44%	Indels:	23
DB:	2	Gaps:	2

US-09-890-220-2 (1-445) x US-08-798-744-25 (1-1430)

```
QY 372 AsnHISgLYLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCys 391
    |||:||||| :|||: ||||| ||| |||||
DB 551 AACATGGCTACTCGAGATTACACACACACACACACATGGCTACTCGAGATTAC 610
QY 392 ArgAsnSerSerAsp----- 396
    |||:|||||:
DB 611 AACACACACACACATGGCTACTCGAGATTACACACACATGGCTACTCGAAAAAATC 670
QY 397 -----ThrThrThrAsnAsnAsnAsnSerValAspArgProSer 410
    |||:|||||:
DB 671 AACAAATATGGTTACTCCAGATTACATGAACACACACACGCTTCTCCGAGAGTTAC 730
QY 411 AspSerAsnThrAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsn 430
    :|||:||||| ||||| ||||| ||| |||
DB 731 AACACCAACACACACACATATTTCTCCGAGATTACACACACATATATACAAAT 790
QY 431 AsnVal-----AspAsnLysAspAsnAsnSerArgAspLysVal 443
    |||:||||| :|||: ||||| ||| |||
DB 791 AATGTTTCTCCGAGATTACACACACATATATACAAATATATGTT 835
```

Search completed: June 20, 2003, 01:25:08
Job time : 95 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 20, 2003, 00:55:28 ; Search time 231 Seconds
(without alignments)
2826.862 Million cell updates/sec

Title: US-09-890-220-2
2378
Perfect score: 1 MCRONCRACKSSPEEVISTDE.....INKNKNVDKDNNSRDYIK 445
Sequence:

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1042519 seqs, 733713590 residues
Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n model -DEV=xlh
-O=/cgn2.1/USFto_spool/US09890220/runat_13062003.144854.16586/app.query.fasta.1.583
-DB=Published_Applications_NA -OFM=fastap -SUFFIX=trpb -MINMATCH=0.1
-DOOPUL=0 -DOOPEXT=0 -UNITS=Dits -START=1 -END=1 -MATRIX=Dlosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pcpt -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09890220.ecgn1.1.156.0runat.13062003.144854.16586
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEBUERY -NEG_SCORES=0 -WAIT -DSPBLCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELXT=7

Database :

1: /cgn2.6/ptodata1/pubna/US07_PUBCOMB.seq:*
2: /cgn2.6/ptodata1/pubna/PCCT_NEM_PUB.seq:*
3: /cgn2.6/ptodata1/pubna/US06_NEM_PUB.seq:*
4: /cgn2.6/ptodata1/pubna/US06_PUBCOMB.seq:*
5: /cgn2.6/ptodata1/pubna/US07_NEM_PUB.seq:*
6: /cgn2.6/ptodata1/pubna/PCCT_PUBCOMB.seq:*
7: /cgn2.6/ptodata1/pubna/US08_NEM_PUB.seq:*
8: /cgn2.6/ptodata1/pubna/US08_PUBCOMB.seq:*
9: /cgn2.6/ptodata1/pubna/US09_NEM_PUB.seq:*
10: /cgn2.6/ptodata1/pubna/US09_PUBCOMB.seq:*
11: /cgn2.6/ptodata1/pubna/US10_NEM_PUB.seq:*
12: /cgn2.6/ptodata1/pubna/US10_PUBCOMB.seq:*
13: /cgn2.6/ptodata1/pubna/US60_NEM_PUB.seq:*
14: /cgn2.6/ptodata1/pubna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	423	17.8	2556	US-09-938-842A-2418
2	266	11.2	888	US-09-938-842A-453
3	225	9.5	3545	US-09-764-864-144
4	221	9.3	2217	US-09-874-162A-6

5	221	9.3	2328	US-09-874-162A-9	Sequence 9, Appl1
6	221	9.3	4409	US-09-874-162A-7	Sequence 7, Appl1
7	221	9.3	4441	US-09-874-162A-4	Sequence 4, Appl1
8	221	9.3	4441	US-09-969-347-169	Sequence 169, App
9	157.5	6.6	1257	US-09-822-830A-428	Sequence 428, App
10	153	6.4	2481	US-09-764-864-124	Sequence 124, App
11	129	5.4	4985	US-10-094-240-10	Sequence 35, Appl
12	128	5.4	2001	US-09-801-368-35	Sequence 131, App
13	127.5	5.4	2886	US-09-801-368-131	Sequence 1, Appl1
14	124	5.2	3931	US-10-006-780-1	Sequence 472, App
15	123	5.2	264	US-09-969-373-472	Sequence 42, Appl
16	122.5	5.2	3996	US-10-087-464-42	Sequence 568, App
17	115	4.8	526	US-09-764-864-568	Sequence 100, App
18	113.5	4.8	32768	US-09-070-927A-100	Sequence 11, Appl
19	109	4.6	2237	US-09-994-485-7	Sequence 1872, App
20	109	4.6	103	US-09-938-842A-1872	Sequence 243, App
21	103	4.3	3150	US-09-938-842A-243	Sequence 13, Appl
22	103	4.3	6868	US-10-228-531-13	Sequence 129, App
23	103	4.3	2713	US-09-981-876-129	Sequence 123, App
24	101.5	4.3	2713	US-09-148-545-129	Sequence 582, App
25	101.5	4.3	2713	US-09-801-368-123	Sequence 81, Appl
26	101	4.2	1605	US-09-801-368-123	Sequence 593, App
27	99.5	4.2	1145	US-09-764-864-582	Sequence 225, App
28	98.5	4.1	4775	US-10-146-477A-81	Sequence 9834, App
29	98.5	4.1	7425	US-10-114-170-212	Sequence 163, App
30	97	4.1	1320	US-10-083-357-593	Sequence 3152, App
31	97	4.1	1680	US-09-801-368-255	Sequence 391, App
32	96.5	4.1	5085	US-10-198-846-9854	Sequence 91, Appl
33	96.5	4.1	6723	US-09-070-927A-163	Sequence 223, App
34	96	4.0	333	US-09-938-842A-3152	Sequence 171, App
35	96	4.0	1956	US-09-351-794A-1	Sequence 10, Appl1
36	96	4.0	3282	US-09-801-368-391	Sequence 11, Appl
37	95	4.0	2098	US-10-286-264-91	Sequence 75, Appl
38	95	4.0	2277	US-09-801-368-223	Sequence 6, Appl1
39	95	4.0	9279	US-09-801-368-171	
40	94.5	4.0	496	US-10-155-533-5	
41	94.5	4.0	951	US-10-155-533-10	
42	94.5	4.0	978	US-10-155-533-11	
43	94.5	4.0	978	US-10-155-533-12	
44	94.5	4.0	1672	US-10-153-666-75	
45	94.5	4.0	1678	US-10-153-533-6	

ALIGNMENTS

RESULT 1
US-09-938-842A-2418
Sequence 2418, Application US/09938842A
Patent No. US20020160378A1
GENERAL INFORMATION:
APPLICANT: Harper, Jeff
APPLICANT: Kreps, Joel
APPLICANT: Wang, Xun
TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
FILE REFERENCE: SCDP1300-3
CURRENT APPLICATION NUMBER: US/09/938, 842A
CURRENT FILING DATE: 2001-08-24
PRIOR APPLICATION NUMBER: US 60/227, 866
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: US 60/264, 647
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/300, 111
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2418
LENGTH: 2556
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2418
Alignment Scores:

Pred. No.: 7,14e-41 Length: 2556
 Score: 423.00 Matches: 101
 Percent Similarity: 40.39% Conservative: 2
 Best Local Similarity: 39.61% Mismatches: 0
 Query Match: 17.79% Indels: 152
 DB: 9 Gaps: 2

US-09-890-220-2 (1-445) x US-09-938-842A-2418 (1-2556)

OY 43 ProserpheleuproacgysleuasntylisllcglAlAllysArglyenarglys --- 61
 DB 2101 CCATCGTTTCTGCCAGATCTTGAACCTACAAATGGGCAAGGCCAAGAAAGAAAGATAT 2160
 OY 61 GCGTTTCTTCTGAAGTGAAGTCCACAGTATGATATGATATCTTACTCTTAATATG 61
 DB 2161 GCGTTTCTTCTGAAGTGAAGTCCACAGTATGATATGATATCTTACTCTTAATATG 2220
 OY 62 -----SerargSer 64
 DB 2221 AAGCTGATGAACATATTTATCTTTGAGATATGACATATGAATGAATGGTCAAGATCT 2280
 OY 65 ThrGlymetValAlPheasnTyrlYsaspCyasnAsnThrLeuGlnLysThrGluVal 84
 DB 2281 ACTGGATGTAGATTTCACATTAAGATTTGTAATTAATACATTAACAAAGAACTGAA--- 2337
 OY 85 ArgGluaspCyserCyserPropheCyserMetLeuCyGlySerPheLysGlyLeuGln 104
 DB 2337 ----- 2337
 OY 105 PheHisLeuasnSerSerHisAspLeupheGluPheGluPheLysLeupheGluGluTyr 124
 DB 2337 ----- 2337
 OY 125 GlnThrValasnValSerValLysLeuasnSerPheLlePheGluGluGluGlySerasp 144
 DB 2337 ----- 2337
 OY 145 AspAspLysPheGluProPheSerLeuCySerLysProArgLysArgGlnArgGly 164
 DB 2337 ----- 2337
 OY 165 GlyArgasnAsnThrArgArgLeuLysValCyPheLeuProLeuAspSerProSerLeu 184
 DB 2337 ----- 2337
 OY 185 ThrAsnGlyThrGluasnGlyLleThrLeuLeuasnAspGlyAsnArgGlyLeuGlyTyr 204
 DB 2338 -----GGAACCGTGGTTTGAAGATAT 2358
 OY 205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnLleProProAlaLle 224
 DB 2359 CCGGAGGCAACAGAGCTGCTGACAAATTGAGACATGACAGCAACATTCGCCAGCCATA 2418
 OY 225 AlaHisSerLeuaspAlaGlyAlaLysValLleLeuThrSerGluAlaValPro 244
 DB 2419 GCCCACTCTTCTGAGACGCTGCTGAAGTTAATTAACCAACCGAAGCTGTGTCCT 2478
 OY 245 AlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
 DB 2479 GCTACTAAGACAGAAAGTTATCTGCTGAGCATGAGAGGCTAGA 2523

RESULT 2
 US-09-938-842A-453
 ; Sequence 453, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Man, Yung
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SCRIP1300-3

; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; CURRENT FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO: 453
 ; LENGTH: 888
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-453

Alignment Scores:
 Pred. No.: 1.93e-22 Length: 888
 Score: 266.00 Matches: 70
 Percent Similarity: 55.73% Conservative: 30
 Best Local Similarity: 53.44% Mismatches: 28
 Query Match: 11.19% Indels: 28
 DB: 9 Gaps: 4

US-09-890-220-2 (1-445) x US-09-938-842A-453 (1-888)

OY 278 MetAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspAsp 297
 DB 1 ATGGCGCTTGAGACAGATATGCTGATCGGATAGCAGATGCAATCGATCGATGTT 60
 OY 298 AlaAspPheGluAspArgGln-----MetLeuAspAspPheValAspAlaAsnLysAspGluGln 304
 DB 61 GCAGATTTTGAAGATCGCCAGCATGATTCATGATTTCTTGCGTTCAATAGTAGGCA 120
 OY 305 -----MetLeuAspAspPheValAspAlaAsnLysAspGluGln 318
 DB 121 CAGAAATGCTATGATATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 OY 319 PheMetHisLeuThrPasnSerPheValArgLysGlnArgValLleAlaAspGlyHisLle 338
 DB 181 TTCATGCAATCTTGGAACTGCTTGTAAAGAAACAAAGA-----AACAAACACATC 231
 OY 339 SerThrPalacyGluAlaPheSerArgPheThrGluLysGluLeuHisArgTyrSerSer 358
 DB 232 TTATTACCAATCTTGTCTTCAATAGCATCCTCTCGACAGCTTACAC-----CTCTCC 285
 OY 359 LeuphetPcystrPargLeupheLeuLleLysLeuThrPasnHisGlyLeuValAspSer 378
 DB 286 TTA-----ACTCTTACCTCTTCTGCTTAAACCAACATGCTTTGTTAGAA 333
 OY 379 AlaThrLleAsnAsnCyasnThrLleLeuGlu 389
 DB 334 GCCTTAGCTAACGTTGTGATACACAGGTTTGA 366

RESULT 3
 US-09-764-864-144
 ; Sequence 144, Application US/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; PRIOR FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1792
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 144
 ; LENGTH: 3545
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-764-864-144

Alignment Scores:

Pred. No.: 1,99e-16 Length: 3545
 Score: 225.00 Matches: 84
 Percent Similarity: 37.09% Conservative: 64
 Best Local Similarity: 21.05% Mismatches: 151
 Query Match: 9.46% Indels: 100
 DB: 10 Gaps: 12

US-09-890-220-2 (1-445) x US-09-764-864-144 (1-3545)

QY 58 ArgLysArgLys-----SerArgSerThrGlyMetValValPheAsnTyr 72
 DB 503 AGAAAGAAAAGAGATCTCCAAATGAAACCGACAAAATTAAGAAATTTTATACACTT 562
 QY 73 LysAspCysAsnThrLeuGlnLysThrGluValArgLysPysCysSerCysProPhe 92
 DB 563 CTCTATACAAACATATACAGCAACAACCTGAACCAAGATGACCTGCAATGCCCTTGG 622
 QY 93 CysSerMetLeuGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
 DB 623 TGACTCTGAACTGCCCAACTTATAGTTTACTCAAGCATCTTAAACTCTGCCATAGC 682
 QY 113 LeuPheGluPheGluPheLysLeuPheGluTyrGlnThrValAsnValSerValLys 132
 DB 683 AGATTATCTTCAACTATGTTTATCATCCAAAGCTGCTAGATAGATGTTTCTATC-- 739
 QY 133 LeuAsnSerPheIlePheGluGluGlySerAspAspLysPheGluProPheSer 152
 DB 740 -----AATGAGGTTATGATGCG-----TCC 760
 QY 153 LeuCysSerLysProArgLys---ArgArgGlnArgGly-----GlyArgAsnAsn 168
 DB 761 TATGACAGAAATCTCAGATATTCATCGCCACCTGATGCTTTAGTCGCAACGGA 820
 QY 169 ---ThrArgArgLeuLysValLysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
 DB 821 CCAAGTTAAGAAACACCTATCACACATATTTCTGTGCGAGCGCAAAAACGACAAAAGCA 880
 QY 188 ThrGluAsnGlyIleThrLeuLeuAsnGlyAsnArgGlyLeuGlyTyrProGluAla 207
 DB 881 AGCATGTCTGAATTTCTGAAATCTGAAGATGCG----- 913
 QY 208 ThrGluLeuAlaGlyInPheGluMetThrSerAsnIleProAlaIleAlaHisSer 227
 DB 913 ----- 913
 QY 228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValProAlaThrLys 247
 DB 913 ----- 913
 QY 248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArg 267
 DB 914 -----GAAGTACAGACGCAAAAGACATATAGTAGTGCCAC-----AATCGT 955
 QY 268 GlnPheTyrHisSerHisArgValGlnProMetAlaLeuGlnGluValMetSerAspArg 287
 DB 956 CTGTATTTCCATGAGATCCGTGCTTACCTCCGCTCCACAAAGAAATG-----GAAGTA 1009
 QY 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
 DB 1010 GATAGTGCAAGATGAAAAGATGCTGAATGCTAAGACAAAACCATTTACCAAAATTGAA 1069
 QY 308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrPAsnSerPheVal 327
 DB 1070 GAGTTTCTGATGTATAGAGAGAGAGAAAGAAAGTGAATCTGGAATCTCCATGTC 1129
 QY 328 ArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArg 347
 DB 1130 ATGAAAGCATGGTTATTTCTGACAAATCAATGATCATGCTGTATGCTTTGTAGAA 1189
 QY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTyrCysThrArgLeuPheLeu 367
 DB 1190 AATTATGACAGAAATTAATTAG--AAGAAATTATGTGCAAACTCATGCTTCATCTA 1246

QY 368 IleLysLeuThrPAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
 DB 1247 GTCAGCATGATGATCACTTATCTTATAGCATATATGCAATAGATTAAGCTGTACCAAG 1306
 QY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
 DB 1307 CTCCTGTAAGATGACG-----CAAAATATGAA 1333
 QY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
 DB 1334 AAGGGGAATCTGCTTCCCTCGCAACGAAATTAAGTAAAGAAATTTGGACACGA 1393
 QY 425 -----AspIleAsnLysAsnValAspAsnLysAspAsnAsnSer 439
 DB 1394 AATGATTTAGTGAATTAATTAAGTAAAGAAAGAAAGCTTTGGAACAGATAGTCTCA 1450

RESULT 4

US-09-874-162A-6
 ; Sequence 6, Application US/09874162A
 ; Patent No. US2002015452A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Koonitz, Jason
 ; APPLICANT: Sklar, Jeffrey
 ; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZF1 GENES IN
 ; TITLE OF INVENTION: ENDOMETRIAL STROMAL TUMORS
 ; FILE REFERENCE: 05311-024001
 ; CURRENT APPLICATION NUMBER: US/09/874,162A
 ; CURRENT FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: US 60/209,093
 ; PRIOR FILING DATE: 2000-06-02
 ; NUMBER OF SEQ ID NOS: 23
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 6
 ; LENGTH: 2217
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-09-874-162A-6

Alignment Scores:

Pred. No.: 2.81e-16 Length: 2217
 Score: 221.00 Matches: 84
 Percent Similarity: 37.09% Conservative: 64
 Best Local Similarity: 21.05% Mismatches: 151
 Query Match: 9.29% Indels: 100
 DB: 9 Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-6 (1-2217)

QY 58 ArgLysArgLys-----SerArgSerThrGlyMetValValPheAsnTyr 72
 DB 1237 AGAAAGAAAAGAGATCTCCAAATGAAACCGACAAAATTAAGAAATTTTATACACTT 1296
 QY 73 LysAspCysAsnThrLeuGlnLysThrGluValArgLysPysCysSerCysProPhe 92
 DB 1297 CTCTATACAAACATATACAGCAACAACCTGAACCAAGATGACCTGATGCTTGG 1356
 QY 93 CysSerMetLeuGlySerPheLysGlyLeuGlnPheHisLeuAsnSerSerHisAsp 112
 DB 1357 TGACTCTGAACTGCCCAACTTATAGTTTACTCAAGATCTTAAACTCTGCCATAGC 1416
 QY 113 LeuPheGluPheGluPheLysLeuPheGluTyrGlnThrValAsnValSerValLys 132
 DB 1417 AGATTATCTTCAACTATGTTTATCATCCAAAGCTGCTAGATAGATGTTTCTATC-- 1473
 QY 133 LeuAsnSerPheIlePheGluGluGlySerAspAspLysPheGluProPheSer 152
 DB 1474 -----AATGACTGTTATGATGCG-----TCC 1494
 QY 153 LeuCysSerLysProArgLys---ArgArgGlnArgGly-----GlyArgAsnAsn 168
 DB 1495 TATGACAGAAATCTCAGATATTCATCGCCACCTGATGCTTTAGTCGCAACGGA 1554
 QY 169 ---ThrArgArgLeuLysValLysPheLeuProLeuAspSerProSerLeuThrAsnGly 187

```

Db 1555 CCAGTTAAGAGAACACCTATCCACATATTTCTGTGTGACAGGCCAAAGCAAAAGCA 1614
      |||||
Qy 158 ThrGluasnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
      |||||
Db 1615 AGCATGCTGAAATTTCTTGAATCTGAAGATGG----- 1647
Qy 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProlaIleAlaHisSer 227
      |||||
Db 1647 ----- 1647
Qy 228 SerLeuAspAlaGlyAlaValIleLeuThrSerGluAlaValProAlaThrIlys 247
      |||||
Db 1647 ----- 1647
Qy 248 ThrArgIleuSerAlaGluArgSerGluAlaArgSerHisIleuLeuGlnIlySarg 267
      |||||
Db 1648 -----GAAGTGAAGACAGCAAGAACATATATAGTGGCCAC-----AATCGT 1689
Qy 268 GlnPheTyrHisSerHisArgValGlnPrometAlaLeuGlnGlnValMetSerAspArg 287
      |||||
Db 1690 CTGATATTCATAGTACCTGCTGCTCCGTCACAGAAATG-----GAAGTA 1743
Qy 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
      |||||
Db 1744 GATAGTGAAATGAAAGATCCGATGCGTAAAGAAACCATTTACACAAATTGAA 1803
Qy 308 AspPheValAspValAsnIlyAspGlyGlnPheMetHisIleuThrAsnSerPheVal 327
      |||||
Db 1804 GAGTTTCTGATGTATTAAGAGAGAGAAAGATGAACTGCGAATCTGCATCTCATGTC 1863
Qy 328 ArgIlyGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGlnAlaPheSerArg 347
      |||||
Db 1864 ATGAAGCATGGGTTATTTCTGACAAATCAATCAATCAATCCGTTGCTGTTGTAGAA 1923
Qy 348 PheTyrGluIlyGlnLeuHisArgTyrSerSerLeuPheThrPcysThrArgLeuPheLeu 367
      |||||
Db 1924 AATTATGACAGAAATTAATTAAG--AGAAATTTATTCGAAATCTCATGCTCATCA 1980
Qy 368 IleIlyLeuThrPasnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
      |||||
Db 1981 GTCAAGCATGATGCTTAATCTTATTAATGCAATATGATTAAGTAACTTTTACCAAG 2040
Qy 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
      |||||
Db 2041 CTCGGTGAATGCA-----CAAAATTTAGAA 2067
Qy 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
      |||||
Db 2068 AAGGGGGAATCTGCTCCCTGCAGAAAGAAATTAAGAAATGGAATGGACAGCA 2127
Qy 425 -----AspIleAsnAsnIlyAsnAsnValAspAsnIlyAspAsnAsnSer 439
      |||||
Db 2128 AATGATTTAGTGAATTAATTAACCAAAAGAAAGAACTTGGAAACAGATGTGTCTCA 2184

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-874-162A-9
Alignment Scores:
Pred. No.: 3,05e-16 Length: 2328
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Gaps: 12
US-09-890-220-2 (1-445) x US-09-874-162A-9 (1-2328)
Qy 58 ArgIlySArgIlyS-----SerArgSerThrGlyMetValIlePheAsnTyr 72
      |||||
Db 1348 AGAAAGAAAGAAAGATCTCAAAATGAAACCAAAATTAAGAAATATTTATTCAGTTT 1407
Qy 73 LysAspCysAsnAsnThrLeuGlnIlySerThrGluValAlaArgGluAspCysSerCysProPhe 92
      |||||
Db 1408 CTCTATTAACAAACATATACAGGACACAAACTGAACCAAGATGACCTGATCCCTGG 1467
Qy 93 CysSerMetLeuCysGlySerPheIlySgIlyLeuGlnPheHisIleuAsnSerHisAsp 112
      |||||
Db 1468 TGTACTCTGAACTGCGCAAACTTTATAGTTTACTCAAGCATCTTAACTGTCGATAC 1527
Qy 113 LeuPheGluPheGluPheIlyLeuPheGluIlyGlyGlnThrValAsnValSerValys 132
      |||||
Db 1528 AGATTATCTCACTATGTTTATTCATCCAAAGCTGCTAGATATGTTTCTATC 1584
Qy 133 LeuAsnSerPheIlePheGluGlnGlySerAspAspIlySargIlyPheGluProPheSer 152
      |||||
Db 1585 -----AATGAGTGTATGATGCG-----TCC 1605
Qy 153 LeuCysSerIlyProArgIlyS-----ArgArgGlnArgIly-----GlyArgAsnAsn 168
      |||||
Db 1606 TATGCAAGAAATCTCAAGATATTCATGTCGCAACCTGATTTTACTGCAACAGCA 1665
Qy 169 ---ThrArgArgLeuIlyValIlyCysPheIleuProLeuAspSerProSerLeuThrAsnGly 187
      |||||
Db 1666 CCAGTTAAGAACACCACTATCAACATATTTCTGTGCGAGGCCAAAGCAAAAGCA 1725
Qy 188 ThrGluAsnGlyIleThrLeuLeuAsnAspIlyAsnArgIlyLeuGlyTyrProGluAla 207
      |||||
Db 1726 AGCATGCTGAAATTTCTTGAATCTGAAGATGG----- 1758
Qy 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProlaIleAlaHisSer 227
      |||||
Db 1758 ----- 1758
Qy 228 SerLeuAspAlaGlyAlaValIleLeuThrSerGluAlaValProAlaThrIlys 247
      |||||
Db 1758 ----- 1758
Qy 248 ThrArgIlyLeuSerAlaGluArgSerGluAlaArgSerHisIleuLeuGlnIlySarg 267
      |||||
Db 1759 -----GAAGTGAAGACAGCAAGAAACATATATAGTGGCCAC-----AATCGT 1800
Qy 268 GlnPheTyrHisSerHisArgValGlnPrometAlaLeuGlnGlnValMetSerAspArg 287
      |||||
Db 1801 CTGATTTTCCATAGATGATACCTGCTTACCTCTCCGTCACAAAGAAATG-----GAAGTA 1854
Qy 288 AspSerGluAspGluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
      |||||
Db 1855 GATAGTGAAGAGAAAGATCTGATATGCGTGAAGAAACCAATTTACACAAATTGAA 1914
Qy 308 AspPheValAspValAsnIlyAspGlyGlnPheMetHisIleuThrAsnSerPheVal 327
      |||||
Db 1915 GAGTTTCTGATGATTAAGAGAGAAATGATGATGAATCTGCAATCTCCATGTC 1974
Qy 328 ArgIlyGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSerArg 347
      |||||
Db 1975 ATGAAGCATGGGTTATTTATTCGACAAATCAATGATATGCTGTGTGTAGAA 2034

```



```

OY 348 PheTYGluLYSGluLeuHISarGTYrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2035 AATTATGGACAGAAATAATTAG--AAGATTATATGCGAAACTTCAGCTTCACTTA 2091
OY 368 IleLYSLeuTrpAsnHISGLYLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2092 GTCAGCATGCATGACTTAACTTATTTAGCATAAAGTCATAGATTAACCTGTACCAAG 2151
OY 388 LeuGLuAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2152 CTCCTGGAATGCAG-----CAAAATTAGAA 2178
OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2179 AAGGGGAATCTGCTTCCCTCGCAACGAGAAATACCTGAGAACAAATGGGACAGCA 2238
OY 425 -----AspIleAsnAsnLYSAsnAsnValAspAsnLYSAspAsnAsnSer 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2239 AATGATTATTAGTAATTAATCAAAAGAAAGCTTTGGAAACAGATAGTGTCTCA 2295

RESULT 6
US-09-874-162A-7
; Sequence 7, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: KOONTZ, JASON
; APPLICANT: SKLAR, JEFFREY
; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZ1 GENES IN
; FILE REFERENCE: 05311-024001
; CURRENT APPLICATION NUMBER: US/09/874,162A
; PRIOR FILING DATE: 2001-06-04
; PRIOR APPLICATION NUMBER: US 60/209,093
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 4409
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (52)...(2379)
US-09-874-162A-7

Alignment Scores:
Pred. No.: 8,83e-16 Length: 4409
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Db: 9 Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-7 (1-4409)
OY 58 ArgLYSArgLYS-----SerArgSerThrGLYMetValAlaPheAsnTYR 72
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1399 AGAAAGAAAGAAAGAAATCTCCAAATGAAACCGCAAAATAATTAGAAATATTATATCAAGTTT 1458
OY 73 LYSAspCysAsnAsnThrLeuGLInuYsThrGLuValArgLYSAspCysSerCysProphe 92
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1459 CTTCTATACACATATCAAGGACACAACTGAAAGCAAGAGATGACCTGATGCCCTTGG 1518
OY 93 CysSerMetLeuCysGLYSerPheLYSGLYLeuGLInPheHISLeuAsnSerSerHisAsp 112
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1519 TGTACTCTGAACTGCGCGCAAACTTATAGTTACTCAACACACTTAACTGTGCCATAGC 1578
OY 113 LeuPheGLuPheGLuPheLYSLeuPheGLuGLInuTYrGLInThrValAsnValSerValLYS 132
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1579 AGATTATTATTCATCACTATGTTTATATCATCCAAAGGCTGTAAGATATGATGTTCTATATC--- 1635
OY 133 LeuAsnSerPheIlePheGLuGLInuGLYSerAspAspAspLYSPheGLuPhePropheSer 152

```

```

Db 1636 -----AATGAGTGTATATGATGCG-----TGC 1656
OY 153 LeuCYSSerLYSProArgLYS---ArgArgGLInArgLY-----GLYArgAsnAsn 168
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1657 TATCGAGAAATCTCAGGATATATTCATGCCAACCGGATTTGCTTTAGTGCACAGCA 1716
OY 169 ---ThrArgArgLeuLYSValCysPheLeuProLeuAspSerProSerLeuThrAsnGLY 187
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1717 CCAGTTAAGAGAACCTATACACATATTTCTGTGTGCAGGCGCAAAACGAAACAAAGCA 1776
OY 188 ThrGLuAsnGLYIleThrLeuLeuAsnAspGLYAsnArgLYLeuGLYTrpProGLuAla 207
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1777 AGCATGCTGAATTTCTTGAATCTGAGATGG----- 1809
OY 208 ThrGLuLeuAlaGLInPheGLuMetThrSerAsnIleProProAlaIleAlaHisSer 227
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1809 ----- 1809
OY 228 SerLeuAspAlaGLYAlaLYSValIleLeuThrSerGLuAlaValAlaProAlaThrLYS 247
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1809 ----- 1809
OY 248 ThrArgLYSLeuSerAlaGLuArgSerGLuAlaArgSerHisLeuLeuGLInuLYSArg 267
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1810 -----GAAGTAGAACAGCAAGAAACATATAGTAGGGCCAC-----AATCGT 1851
OY 268 GLInPheTYrHisSerHisArgValGLInProMetAlaLeuGLInValMetSerAspArg 287
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1852 CTGTATTTCCATAGTGAATGATCCCTGCTTACCTCCGCTCCGACAGAAATG-----GAAGTA 1905
OY 288 AspSerGLuAspGLuValAlaAspAspAspValAlaAspPheGLuAspArgGLInMetLeuAsp 307
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1906 GATAGGAGATGAGAAAGATCTGATGCTGATAGAGAAACCACTTACACAAATGGA 1965
OY 308 AspPheValAspValAsnLYSAspGLYLYSGInPheMetHISLeuTrpAsnSerPheVal 327
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1966 GAGTTTTCGATGTTAATGATGAGAGAAAGATGATGAACCTGGAATCTCCATGTC 2025
OY 328 ArgLYSGInArgValIleAlaAspGLYHISLeSerTrpAlaCysGLuAlaPheSerArg 347
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2026 ATGAAAGCATGGTTATGTTCTGTCGACATCAAAATCAATCATGCTGTGCTGTAGAA 2085
OY 348 PheTYGluLYSGluLeuHISarGTYrSerSerLeuPheTrpCysTrpArgLeuPheLeu 367
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2086 AATTATGACAGAAATAATTAG--AAGATTATATGCGAAACTTCAGCTTCACTTA 2142
OY 368 IleLYSLeuTrpAsnHISGLYLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2143 GTCAGCATGCATGACTTAACTTATTTAGCATAAATGATCAATAGATTAAGCTGTACCAAG 2202
OY 388 LeuGLuAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerValAsp 407
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2203 CTCCTGGAATGCAG-----CAAAATTAGAA 2229
OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2230 AAGGGGAATCTGCTTCCCTCGCAACGAGAAATACCTGAGAACAAATGGGACAGCA 2289
OY 425 -----AspIleAsnAsnLYSAsnAsnValAspAsnLYSAspAsnAsnSer 439
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 2290 AATGATTATTAGTAATTAATCAAAAGAAAGCTTTGGAAACAGATAGTGTCTCA 2346

RESULT 7
US-09-874-162A-4
; Sequence 4, Application US/09874162A
; Patent No. US20020155452A1
; GENERAL INFORMATION:
; APPLICANT: KOONTZ, JASON
; APPLICANT: SKLAR, JEFFREY
; TITLE OF INVENTION: FUSION OF JAZF1 AND JAZ1 GENES IN
; FILE REFERENCE: 05311-024001

```

```

CURRENT APPLICATION NUMBER: US/09/874,162A
CURRENT FILING DATE: 2001-06-04
PRIOR APPLICATION NUMBER: US 60/209,093
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 23
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
LENGTH: 4441
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (195)...(2411)
US-09-874-162A-4

Alignment Scores:
Pred. No.: 8,94e-16 Length: 4441
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Gaps: 12

US-09-890-220-2 (1-445) x US-09-874-162A-4 (1-4441)

OY 58 ArglySerThrglyMetValPheAsnTyr 72
Db 1431 AGAAAAGAAAAGGATCTCCAAATGAAAACCGCAAAATTAAGAAATTTATTCAGTTT 1490

OY 73 LysAspCysAsnAnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
Db 1491 CTCATATACAAACATACAGCAAGCAACAACTGAGCAAGAGATGACCTGATGCCCTTGG 1550

OY 93 CysSerMetLeuGlySerPheLysGlyLeuGlnPheHisLeuAsnSerHisAsp 112
Db 1551 TGTACTGTGAACCTCCGCAAACTTTATGTTTTCACAGCACTTAATCTGCCATAGC 1610

OY 113 LeuPheGluPheGluPheLysLeuPheGluGluTyrGlnThrValAsnValSerValLys 132
Db 1611 AGATTATCTTCAACTATGTTTATCTCCAAAGGCTGAGATGATGTTTCTATTC- 1667

OY 133 LeuAsnSerPheLeuPheGluGluGlySerAspAspAspLysPheGluProPheSer 152
Db 1668 -AATGAGTGTATGATGGC-----TCC 1688

OY 153 LeuGlySerLysProArgLys--ArgArgGlnArgGly-----GlyArgAsnAsn 168
Db 1689 TATGCAAGGAATCTCAGATATTCATCCCAACCTGGATTTGCTTTAGTCGCAAGCGA 1748

OY 169 ---ThrArgArgLeuLysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187
Db 1749 CCAAGTAAAGAAACACCTATTCACACATATTTCTGTGCGCAAGCCCAAAAGCA 1808

OY 188 ThrGluAsnGlyLeuThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyTyrProGluAla 207
Db 1809 AGCATGTCTGAATTTCTGAATCTGAATCGAATGG----- 1841

OY 208 ThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProProAlaIleAlaHisSer 227
Db 1841 ----- 1841

OY 228 SerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaValAlaProAlaThrLys 247
Db 1841 ----- 1841

OY 248 ThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLysArg 267
Db 1842 -----GAAGTAAAGCAAGCAAAACATATTAAGTATGCGCAC-----AAACGCT 1883

OY 268 GlnPheThrHisSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAspArg 287
Db 1884 CTGTATTTCCATGATGATACCTGCTTACCTCTCCCAAGAAATG-----GAAGTA 1937

```

```

OY 288 AspSerGluAspGluValAlaAspAspValAlaAspPheGluAspArgGlnMetLeuAsp 307
Db 1938 GATAGTGAAGATGAAGAAAGGATCTGAAATGGCTTAAGAGAAAACCAATTACAAATTTGAA 1997

OY 308 AspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPheVal 327
Db 1998 GAGTTTTCATGATTTAAAGAGAGAGAAAGTGAAGTGAACCTGGAATTTCCATGTC 2057

OY 328 ArgLysGlnArgValIleAlaAspGlyHisLeuSerThrAlaCysGluAlaIlePheSerArg 347
Db 2058 ATGAAAGCATGGGTTTATTTGACAAATCAATGAATCATGCTGTATGCTGTTTGTGAA 2117

OY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheThrPcysThrArgLeuPheLeu 367
Db 2118 AATTATGACACGAAATATATTAAG---AAGAAATTTATGTGAAACCTTCATGCTCAATCTA 2174

OY 368 IleLysLeuThrPasnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle 387
Db 2175 GTACACATGACATGACTTTAATCTTATAGCATATGATGATGATGATGATGATGATGATG 2234

OY 388 LeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerValAsp 407
Db 2235 CTCGCTGAATGACAG-----CAAAATTTAGAA 2261

OY 408 ArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisProAsn----- 424
Db 2262 AAGGGGGAATGCTGCTCCCGCAAGCAAGAAATATACGAAACAAATGGGACAGCA 2321

OY 425 -----AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSer 439
Db 2322 AATGATTTTATGTAATTAATCTCAAAAGAGAAACCTTTGGAACAAGATGCTCTCA 2378

RESULT 8
US-09-969-347-169
Sequence 169, Application US/09969347
Patent No. US20020115085A1
GENERAL INFORMATION:
APPLICANT: Eber, Reinhard
TITLE OF INVENTION: Sets
FILE REFERENCE: 689290-69
CURRENT APPLICATION NUMBER: US/09/969, 347
PRIOR FILING DATE: 2000-10-02
CURRENT FILING DATE: 2001-10-02
PRIOR APPLICATION NUMBER: US/60/237, 598
PRIOR FILING DATE: 2000-10-03
NUMBER OF SEQ ID NOS: 318
SOFTWARE: PatentIn version 3.0
SEQ ID NO 169
LENGTH: 4441
TYPE: DNA
ORGANISM: Homo sapiens
US-09-969-347-169

Alignment Scores:
Pred. No.: 8,94e-16 Length: 4441
Score: 221.00 Matches: 84
Percent Similarity: 37.09% Conservative: 64
Best Local Similarity: 21.05% Mismatches: 151
Query Match: 9.29% Indels: 100
Gaps: 12

US-09-890-220-2 (1-445) x US-09-969-347-169 (1-4441)

OY 58 ArglySerThrglyMetValPheAsnTyr 72
Db 1431 AGAAAAGAAAAGGATCTCCAAATGAAAACCGCAAAATTAAGAAATTTATTCAGTTT 1490

OY 73 LysAspCysAsnAnThrLeuGlnLysThrGluValArgGluAspCysSerCysProPhe 92
Db 1491 CTCATATACAAACATACAGCAAGCAACAACTGAGCAAGAGATGACCTGATGCCCTTGG 1550

```

QY	93	CysSerMetLeuGlySerGlySerPheLeuGlyLeuGlnPheHisLeuAsnSerSerHisasp	112
Db	1551	TTGACTCTGAACCTCCCCCAACTTTAATAGTTTACTCAAGCATCTTAACTCTGCATACGC	1610
QY	113	LeuPheGlnupheGlnupheLeuPheGlnupheGlnupheGlnupheValAsnValSerVallys	132
Db	1611	AGATTTATTCCTCAACSTATGTTTATTCATCCAAAGAGTGTAGCATAGATGTTTTCATC---	1667
QY	133	LeuAsnSerPheIlePheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnuphe	152
Db	1668	-----AATGACGTGTTATCATGAGC-----TTC	1688
QY	153	LeuGlySerLeuPheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnuphe	168
Db	1689	TATCCAGAGAAATCTCTCAGATATTCATCCGCCAACCTGTGATTTGCTTTAGTCGCAACGGA	1748
QY	169	---ThrArgArgLeuValCysPheLeuPheLeuAsnSerPheSerLeuThrAsnGly	187
Db	1749	CCAGTTAAGAGAACCCCTATCATCCAAATATCTGTGTGTCAGCGCCAAACGAAACGA	1808
QY	188	ThrGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnuphe	207
Db	1809	AGCATGTCTGATTTCTTGAATCTGAAGATGG-----	1841
QY	208	ThrGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnuphe	227
Db	1841	-----	1841
QY	228	SerLeuAsnAlaGlyAlaValIleLeuThrSerGlnAlaValAlaProAlaThrLys	247
Db	1841	-----	1841
QY	248	ThrArgLysLeuSerAlaGluArgSerGlnAlaArgSerHisLeuLeuGlnLysArg	267
Db	1842	-----GAGTGAAGAACAGCAAAAGACATATGATGAGTGGCCAC-----AATCGT	1883
QY	268	GlnPheThrHisSerHisArgValGlnProMetAlaLeuGlnGlnValMetSerAspArg	287
Db	1884	CTGATATTTCCATAGCTGATACCTGCTGCTCTCCGTCACCAAGAAATG-----GAAGTA	1937
QY	288	AspSerGlnAspGlnValAspAspAspValAlaAspPheGlnupheGlnupheGlnupheGlnuphe	307
Db	1938	GATAGTGAAGATGAAAAAGATCTGCAATGCGTAAAGACAAAAACCATTACCAAAATGGA	1997
QY	308	AspPheValAspValAsnLysAspGlnLysGlnPheMetHisLeuThrAsnSerPheVal	327
Db	1998	GAGTTTCTGCTGTTATTAAGAGAGAGAAACAGACGATGAAGAACTGTGAATCTCATGTC	2057
QY	328	ArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGlnAlaPheSerArg	347
Db	2058	ATGAGAGATGGGTTATGCTGACAAATCAATGAATCATGCTGTATGCTTTGTAGAA	2117
QY	348	PheArgGlnGlnLysGlnLeuHisArgLysSerSerLeuPheThrCysThrArgLeuPheLeu	367
Db	2118	AATATATGACAGAAATAATTAAG---AAGAAATTTATGTCGAAACTTCATGCTTCATCTA	2174
QY	368	IleLeuLeuThrAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIle	387
Db	2175	GTCAGCATGCATGACTTTAATCTTAATAGCATTAATGTCMAATAGATAAGCTGTATCCAA	2234
QY	388	LeuGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnupheGlnuphe	407
Db	2235	CTCCGCTGAATGCAG-----CAAAAATTGAA	2261
QY	408	ArgProSerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsn-----	424
Db	2262	AAGGGGAATCTGCTCCCTCGCAAGAGAAATTAATCTGAAGAAACAAATAGGACAGACA	2321
QY	425	-----AspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnAsnSer	439
Db	2322	AATGATTTTACTGAAATTAATCTCAAAAGAGAAACCTTTGGAACAGATAGTGTCTCA	2378

```

US-09-822-830A-428
; Sequence 428, Application US/09622830A
; Patent No. US20020142952A1
; GENERAL INFORMATION:
; APPLICANT: Genetics Institute, Inc.
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J.
; APPLICANT: Howes, Steven H.
; APPLICANT: Resnick, Richard J.
; APPLICANT: Gulikota, Kamalakar
; APPLICANT: Graham, James R.
; TITLE OF INVENTION: POLYNUCLEOTIDES ENCODING NOVEL SECRETED PROTEINS
; FILE REFERENCE: GIN 6402
; CURRENT APPLICATION NUMBER: US/09/822,830A
; CURRENT FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/195,604
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 631
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 428
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 860
; OTHER INFORMATION: n=a,c,g, or t
US-09-822-830A-428

```

Alignment Scores:

Pred. No.:	6.36e-09	length:	125
Score:	157.50	Matches:	54
Percent Similarity:	45.658	Conservative:	51
Best local Similarity:	23.488	Mismatches:	103
Query Match:	6.628	Indels:	23
DB:	10	Gaps:	5

US-09-890-220-2 (1-445) x US-09-822-830A-428 (1-1257)

Oy		217	TheSerAmLeprProAlaIleLanHisSerSerLeuAspArgValValIle	238
Dd		146	ACGACAGATTAAAGAGACAACCTATCACATATTCTGTGGCGCCA-AAACGAACA	204
Oy		237	LeuThrSerGluAlaValValProAlaThrIlyThraGlysLeuSerAlaGluArgSer	256
Dd		205	AAMGCAAGCATTCGGTAATTTCTGAATCTGAAGAATGGGGGAATGGAACACGAAGAACA	264
Oy		257	GluAlaArgSerHisLeuLeuLeuGlnIlyAspGlnPheTygHisSerHisArgValGln	276
Dd		265	TATAGTAGTGGCCAC-----AATCCTCGTATTTCCACATGATACCTGCTTA	312
Oy		277	ProMetAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAsp	296
Dd		313	CCTCTCCGTCACAAATAATG-----GAACTAGACTAGAAATGAAAAGATCCTGAA	366
Oy		297	ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnIlyAspGlu	316
Dd		367	TGGCTAAGAGAAAAAACCATTCACAAATAATGAAGAGTTTTCTGTTAATGAAGAGAG	426
Oy		317	LysGlnPheMetHisLeuTrpAsnSerPheValArgGlyGlnArgValHisLeaAspGly	336
Dd		427	AAAGAAAGTGAAACTCTGGAAATCTCCATGCATCAAGACATGGGTTATTGCTGCACAAT	486
Oy		337	HisLeuSerTrpAlaGlyGluAlaPheSerArgPheTygIlyGlyGluLeuHisArgTyr	356
Dd		487	CAATGATCATCATCCGTATGCTGTTTGTGAAGAAATATATGGACAGAAATAATATTAAG--	543
Oy		357	SerSerLeuPheTrpCystTrpArgLeuPheLeuIleLysLeuThrPasnHisGlyLeuVal	376
Dd		544	AAGAAATTTGTGTGAAACTTCATGGCTTATCATAGACACATGCACTTAATCTTAATTT	603

QY 377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeu----- 388
 DB 604 AGCATATGTCATTAAGTAAAGCTGTTACCACTCCGTAATATGCAGCAAAATTAGAA 663
 QY 389 -----GluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerVal 406
 DB 664 AAGGGGAAATCTGCTCCCTCCGCAAAAGCAAAATTAATCAAGAACAAATGCGACACA 723
 QY 407 AsparGProSerAspSerAsnThr-----AsnAsnAsnIleValAsp 421
 DB 724 AATGATTTACTGTAATTAATTAACCAAGAACAAAGCTTTGGAAACAGATAGTGTCAAGG 783
 QY 422 HisProAsnAspIleAsnAsnLysAsnAsp 431
 DB 784 TTTCAAAAACAGACCAAAAAACAACTCT 813

RESULT 10

US-09-764-864-124
 ; Sequence 124, Application us/09764864
 ; Patent No. US20020132753A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT23
 ; CURRENT APPLICATION NUMBER: US/09/764,864
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1782
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 124
 ; LENGTH: 2481
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-764-864-124

Alignment Scores:

Pred. No.: 7e-08 Length: 2481
 Score: 153.00 Matches: 53
 Percent Similarity: 44.35% Conservative: 49
 Best Local Similarity: 23.04% Mismatches: 103
 Query Match: 6.43% Indels: 26
 DB: 10 Gaps: 5

US-09-890-220-2 (1-445) x US-09-764-864-124 (1-2481)

QY 217 ThrSerAsnIleProProlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIle 236
 DB 200 ACGGACGAGTTAAGAGAACACCTATCACACATATCTGTGTGCGAGGCA-AAAAGAACCA 258
 QY 237 LeuThrSerGluAlaValProAlaThrIlyThrArgLysLeuSerAlaGluArgSer 256
 DB 259 AAAGCAGCAGTCGTGAATTTCTTGAATCTGAAGGGAAGTGAACACCAACCAACCA 318
 QY 257 GluAlaArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln 276
 DB 319 TATAGTAGTGCCAC-----AATGCTGTATTTTCATAGGATACCTGCTTA 366
 QY 277 ProMetAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAsp 296
 DB 367 CCTCTCGCTCCACAGAAATG-----GAAGTAGATAGTGAAGAGAAAGAGCTTGA 420
 QY 297 ValAlaAspPheGluAspArgIleMetLeuAspAspPheValAspValAsnLysAspGlu 316
 DB 421 TGGCTAAGAGAAAAACCATTAACAATAATTGAAGAGTTTTCGATGTTAATGAAGAGAG 480
 QY 317 LysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
 DB 481 AAAGAACTGATGAACCTGCGAATCTCCATGTCATGAGACAGCATGGTTATATGCTGACAT 540
 QY 337 HisIleSerThrAlaGluAlaPheSerArgPheTyrGluLysGlnIleAsnArgTyr 356
 DB 541 CAATGAAATCATGCTGTATAGCTTTGTAGAAAAATTATAGACAGAAAAATGATTAAG--- 597

QY 357 SerSerIlePheTrpCysTrpArgLeuPheIleLysLeuThrPasnHisGlyLeuVal 376
 DB 598 AAGAAATTTTGTGGAAGCTCATGCTCATCTCATCTGATGATGATGATGATGATGAT 657
 QY 377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSer 396
 DB 658 AGCATATGTCATTAAGTAAAGCTGTTACCACTCCGTAATATGCAGCAAAATTAGAA 705
 QY 397 ThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsn 416
 DB 706 -----CAAAATTTGAAGAGGGAATCTGCTCCCTCCGCAAC 744
 QY 417 AsnAsnIleValAspHisProAsn-----AspIleAsnAsnLys 429
 DB 745 GAAGAAATTAATCAAGAACCAAAATGGGACACCAATGATTTAGTGAATTAATCAATA 804
 QY 430 AsnAsnValAspAsnLysAspAsnAsnSer 439
 DB 805 GAGAAAGCTTTGGAATCAAGATAGTGTCTCA 834

RESULT 11

US-10-094-240-10/C
 ; Sequence 10, Application US/10094240
 ; Publication No. US20030082637A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ZWIEBEL, LAURENCE J.
 ; TITLE OF INVENTION: ARRESTIN GENE, POLYPEPTIDE, AND METHODS OF USE THEREOF
 ; FILE REFERENCE: N8289
 ; CURRENT APPLICATION NUMBER: US/10/094,240
 ; CURRENT FILING DATE: 2001-03-08
 ; PRIOR APPLICATION NUMBER: 10/056,405
 ; PRIOR FILING DATE: 2002-01-24
 ; PRIOR APPLICATION NUMBER: 60/264,649
 ; PRIOR FILING DATE: 2001-01-26
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 10
 ; LENGTH: 4985
 ; TYPE: DNA
 ; ORGANISM: Anopheles gambiae
 US-10-094-240-10

Alignment Scores:

Pred. No.: 0.000192 Length: 4985
 Score: 129.00 Matches: 33
 Percent Similarity: 52.38% Conservative: 22
 Best Local Similarity: 31.43% Mismatches: 32
 Query Match: 5.42% Indels: 18
 DB: 9 Gaps: 4

US-09-890-220-2 (1-445) x US-10-094-240-10 (1-4985)

QY 348 PheTyrGluLysGluLeuHisArgTyrSerSerLeuPheTrpCysTrpArgLeuPhe 367
 DB 3174 TTCTTAAGATTAATGTTGTTATTAATTAATTAATTAATTAATTAATTAATTAAT 3121
 QY 368 Ile-----LysLeuThrPasnHisGlyLeuVal-----Asp 377
 DB 3120 ATTATTTACTTTATTAATTTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3061
 QY 378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
 DB 3060 ACAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 3019
 QY 398 ThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsn 417
 DB 3018 AATAAGAACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 2959
 QY 418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAsnLysAspAsn 437
 DB 2958 AATATATATACATAGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2899
 QY 438 AsnSerArgAspLys 442

DB 2898 AATAATAAAAAA 2884

RESULT 12

US-09-801-368-35

Sequence 35, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Call, Brian

APPLICANT: Hecht, Peter

APPLICANT: Holtzman, Doug

APPLICANT: Madden, Kevin

APPLICANT: Maxon, Mary

APPLICANT: Milne, Todd

APPLICANT: No. US20020128250A1man, Thea

APPLICANT: Royer, John

APPLICANT: Salama, Sofie

APPLICANT: Sherman, Amir

APPLICANT: Silva, Jeff

APPLICANT: Summers, Eric

TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi

FILE REFERENCE: 109272.147

CURRENT APPLICATION NUMBER: US/09/801,368

PRIOR FILING DATE: 2001-03-07

PRIOR APPLICATION NUMBER: US 09/487,558

PRIOR FILING DATE: 2000-01-19

PRIOR APPLICATION NUMBER: US 60/160,587

NUMBER OF SEQ ID NOS: 440

SOFTWARE: PatentIn version 3.0

SEQ ID NO 35

LENGTH: 2001

TYPE: DNA

ORGANISM: Saccharomyces cerevisiae

US-09-801-368-35

Alignment Scores:

Pred. No.:	5,58e-05	Length:	2001
Score:	128.00	Matches:	86
Percent Similarity:	34.55%	Conservative:	56
Best Local Similarity:	20.92%	Mismatches:	118
Query Match:	5.38%	Indels:	152
	10	Gaps:	19

US-09-890-220-2 (1-445) x US-09-801-368-35 (1-2001)

QY 118 PheIysLeuphegluIutrglnhrValAsnValSerValIysLeuAsnSerPheIle 137

DB 109 TATGAAATAGAAAGACCTTATTGATCGTATCAGTAGTATA-----TTCGTT 156

QY 138 PhegluIugluIuSer-----AspAspAspIysPhegluIuPhe 151

DB 157 TTTATTGAAAGATATATCGGCGCATCAAGAGATGAGTGGCGTTCTTCTGCTCATCA 216

QY 152 SerIeuCySerIys-----ProArgIysArgIgluIuArgIgluIuArgIgluIu 167

DB 217 AGAATATCAGGCAAGTTCTTAATATACAGAACTAGACAGAAAGACCGGCTGTANT 276

QY 168 AsnThrArgArgLeuIysValCysPheLeuProLeuAspSerProSerLeuThrAsnGly 187

DB 277 -----GCTAATGCT 285

QY 188 ThrGluAsnGlyIleThr-----LeuLeuAsnAspGlyAsnArgIgluIu-----Gly 203

DB 286 ACTTCAGTGGCAGTACTGATTCGCCGCGCATTCAGTAGGAGACTTCTGCGCGAGAAAT 345

QY 204 TyrProGluAlaThrGluLeuAla-----GlyIuPhegluIuMetThr 217

DB 346 AACCCCTTCGCGAAATCAAGTACTCCTTGAAGAACACCAAGTTTGTAT----- 399

QY 218 SerAsnIleProProAlaIleAlaHisSerSerLeuAspAla----- 231

DB 400 -----TTCGCACCTACTATGGCCATTCAGTTTGTAGTCAGAGCAACACCTCCATT 453

QY 232 -----GlyIuAluValIleLeuThrSerGluAlaValProAlaThrIys 247

DB 454 TCTCATCAAAACCGTTCGAACTTACCTTTAAAGTATATACGCTGCGTGAAGAAACATA 513

QY 248 ThrArgIysSerIeAluArgSerGluAlaArgSerHisLeuLeuGlnIysArg 267

DB 514 TCTGTCAAGCTG-----AAGACA 531

QY 268 GlnPheThrHisSerHisArgValGlnProMetAlaLeuGluIuAluMetSerArg 287

DB 532 CTCATTCACATCG-----ATTGAAATTTACATATGTCTCTTATTTATTCGTAAAA 585

QY 288 AspSerGluAspIuValAspAspAspValAlaAspPheGluAspArgIuMetLeuAsp 307

DB 586 GATAT-AAAACAAAATGTCTAGTACTCCTAAAGCTTCCTTTTGAAGATGTAG 644

QY 308 AspPheValAspValAsn-----LysAspGluIysGln 318

DB 645 ACCTTCACAGACCTAATTCCTCCGATGGGAATACACCTTAGCAACGTGAAAAACA 704

QY 319 PheMetHisLeuThrPasnSerPheValArgIysGlnArgValIleAlaAspGlyHisIle 338

DB 705 TTCACCTAC-----CACCGG 719

QY 339 SerTrpAla-----CysGluAlaPheSerArg 347

DB 720 TAATGTCCTAATATATTAACAATAAAGCAACTCTTCACCTCCCTGATCTGTAT 779

QY 348 PheTyrGluIysGluLeuHisArgTyrSer-----SerLeuPheTrpCys 362

DB 780 TTCACAGATTAACAATAGTCTAATATTAATGTCTGCGAGATCAATTTCTAGTGC 839

QY 363 TrpArgLeuPheLeuIle-----LysLeuTrp-----AsnHisGlyLeuVal 376

DB 840 AAACAAAATTTATTTATTAATAAATGATGAAAGCTCTGATACCAATCACAATATTCG 899

QY 377 AspSerAlaThrIleAsnAsn----- 383

DB 900 TCCCGCATTCACATCAACAACCTTGTGATGATACGGCAATCCACCTTATATCACTCAATC 959

QY 384 -----Cys-AsnThrIleLeuGluAsnCy 391

DB 960 TCCTGATTAACACTAAGCTACTGTATGATACACATGTTAATATTAATATATATATAG 1019

QY 391 sArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAs 411

DB 1020 TATTAACAGTAAATAGTAATATATAGTAACAATATATAT----- 1059

QY 411 pSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnIysAsnAs 431

DB 1060 -----AAT 1115

QY 431 nValAspAsnIysAspAsnAsnSerArgAsp 441

DB 1116 TGTCAATACTAATATCAGGAGACGTAAACAT 1146

RESULT 13

US-09-801-368-131

Sequence 131, Application US/09801368

Patent No. US20020128250A1

GENERAL INFORMATION:

APPLICANT: Busby, Robert

APPLICANT: Call, Brian

APPLICANT: Hecht, Peter

APPLICANT: Holtzman, Doug

APPLICANT: Madden, Kevin

APPLICANT: Maxon, Mary

APPLICANT: Milne, Todd

APPLICANT: No. US20020128250A1man, Thea

APPLICANT: Royer, John

APPLICANT: Salama, Sofie

APPLICANT: Sherman, Amir
 APPLICANT: Silva, Jeff
 APPLICANT: Summers, Eric
 TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
 FILE REFERENCE: 109272.147
 CURRENT APPLICATION NUMBER: US/09/801,368
 PRIOR FILING DATE: 2001-03-07
 PRIOR APPLICATION NUMBER: US 09/487,558
 PRIOR FILING DATE: 2000-01-19
 PRIOR APPLICATION NUMBER: US 60/160,587
 PRIOR FILING DATE: 1999-10-20
 NUMBER OF SEQ ID NOS: 440
 SOFTWARE: PatentIn version 3.0
 SEQ ID NO 131
 LENGTH: 2886
 TYPE: DNA
 ORGANISM: Saccharomyces cerevisiae
 US-09-801-368-131

Alignment Scores:
 Pred. No.: 0.000118 Length: 2886
 Score: 127.50 Matches: 26
 Percent Similarity: 61.90% Conservative: 13
 Best Local Similarity: 41.27% Mismatches: 23
 Query Match: 5.36% Indels: 1
 Gaps: 1

US-09-890-220-2 (1-445) x US-09-801-368-131 (1-2886)

QY 382 AsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsn 401
 DB 1501 AACCAACACATTAACGATTAACGATTAACATTAATTAATTAATTAACAT 1560
 QY 402 AsnAsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleValAsp 421
 DB 1561 AATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1620
 QY 422 HisProAsnAspIleAsnAsn--LysAsnAsnValAspAsnLysAspAsnAsnSerArg 440
 DB 1621 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1680
 QY 441 AspLysVal 443
 DB 1681 GATTAATTAAT 1689
 RESULT 14
 US-10-006-780-1
 ; Sequence 1, Application US/10006780
 ; Publication No. US20030104496A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sakowitz, Roman
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Guo, Jun
 ; APPLICANT: Freedman, Richard
 ; TITLE OF INVENTION: NOVEL MOTOR PROTEIN OF P. FALCIPARUM AND
 ; FILE REFERENCE: CYTO083
 ; CURRENT APPLICATION NUMBER: US/10/006,780
 ; CURRENT FILING DATE: 2001-11-30
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 1
 ; LENGTH: 3931
 ; TYPE: DNA
 ; ORGANISM: Plasmodium falciparum
 US-10-006-780-1

Alignment Scores:
 Pred. No.: 0.000529 Length: 3931
 Score: 124.00 Matches: 96
 Percent Similarity: 36.36% Conservative: 68
 Best Local Similarity: 21.29% Mismatches: 166
 Query Match: 5.21% Indels: 122

DB: 9 Gaps: 21
 US-09-890-220-2 (1-445) x US-10-006-780-1 (1-3931)
 QY 24 ILeYrCysLysProValArgLeuThrAsnIlePheHisLeuArgSerLeuGluAsnPro 43
 DB 483 ATTTATTTGGT-----AAATTAATGATTTAATTAACAAACGATAGTTG----- 527
 QY 44 SerPheLeuProArgCysLeuAsnTyrIleGluValAspArgLysSerArg 63
 DB 528 -----GTACACCACTT-AGAAAAATGGCAAAAGCA 556
 QY 64 SerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnTyrThrGlu 83
 DB 557 AGTTGTAGTAAAGATTAAATTAATTAAGAGTATTAAACAA-----AGACGAATTAAT 610
 QY 84 ValArgGluAspCysSerCysProPheCysSerMetLeuCys-GlySerPheLysGlyLe 103
 DB 611 ATTAATAATGATAGATGCTTTTATTAAAGAAAAATGGCTTAATTCACAAACGAT-- 668
 QY 103 uGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPheGlu 123
 DB 669 -----GAATCATCTATCATCATGCTATTTAAATGATTTAAAA----- 710
 QY 123 uTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePhe----- 138
 DB 711 -----GATTAATAATAATAATAATACATCTCTGGAAAAATGCTTTCATGATTAGCAGG 763
 QY 139 -GluGluGluGlySerAspAspAspLysPheGluProPheSerLeuCysSerLysProAr 158
 DB 764 AAGTGAAGAGGAGCTGAT-----ACCGTTTCACAAA 796
 QY 158 gLysArgArgGlnArgGlyLysArgAsnThrArg-----ArgLeuLysValCys 175
 DB 797 TAAACAACACCAACCGATGAGCTAATTAATTAATTAATTAATTAATTAATTAATTAAT 856
 QY 175 sPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLe 195
 DB 857 TATTCAGAGTATGAGTTCA-----GATTAATAATAATAATAATACCATTCCTTCAG 898
 QY 195 uAsnAspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu----- 210
 DB 899 AGATTCA-----GATTTAACCTAAAGTTTAAGACATTAAT 934
 QY 211 -AlaGlyGlnPheGlu-----MetThrSerAsnIleProAlaIleAla----- 225
 DB 935 TGTAGGAATCTTAATAAGTATTAATGATAGATCAATTAATTTCTTCAATTAATGATGTA 994
 QY 226 -----HisSerSerLeuAspAlaGlyAlaLysValIleLeuThr 238
 DB 995 GCAAACATTGAATACATTAGATTTCTTCAAGAGTTAAGACCTTAATAATAATAATCTAC 1054
 QY 238 rSerGluAlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGlu 258
 DB 1055 ATGTTAATTAAGTAAAGATGTATACAAATACCGAAAGATTTAGTATTTAGATTCAAG-- 1112
 QY 258 ArgSerHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProHe 278
 DB 1113 -----GCAAGCAAAATGAATGATC 1132
 QY 278 tAlaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAl 298
 DB 1133 TGTATTAAGAAATGATGTT----- 1151
 QY 298 AspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysG 318
 DB 1152 ----ATCAAAATCAATCATCTCTTCAAT-----AATTAATAATAATAATAAT 1195
 QY 318 nPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisI 338
 DB 1196 CATGCTGGCAAAATTAATGATTAAGATGAAGAAAGAAATTAATTAATTAATTAATTAAT 1255
 QY 338 eSerThrPheLysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSe 358

Search completed: June 20, 2003, 02:35:54
Job time : 269 secs

```

DB      1256 TGATTAACCTAGGAGGATTCCACATCAACCTTGGAAAA-----TACAGTTC 1303
QY      358 rleuphetrrpcystpargleupheleuileysleuptrpashisglyleuvalasp-- 377
DB      1304 ACTTAATGATATAGACAAAATAAGAAAAATTAAG-----AAAAAGGTTAATTAATTA 1357
QY      378 -SerAlaThrIleAsnAsnCysAsnThrIleuGlu-----AsnCysArgAs 393
DB      1358 TAAAGTCAATATATATATATGACACACCAATAATAAAGCACACATATATATATTA 1417
QY      393 nSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAs 413
DB      1418 TAATAATGATATATATGATTAATAATATGATATATATATATATATATATATATAGTAG 1477
QY      413 nThrAsnAsnAsnAsnIleValAspHis-----ProAsnAspIleAsnAsnLysAsnAs 431
DB      1478 TAGTATGCTCAATTAACATGATCAATCATATGATCAATTAATTAATTAACACAAATATCA 1537
QY      431 nValAspAsnLysAspAsnAsnSerArgAsp 441
DB      1538 TGTAAATTAACACAAATATATATATATATAT 1568

```

RESULT 15

```

US-09-969-373-472/c
; Sequence 472, Application US/09969373
; Patent No. US20020133852A1
; GENERAL INFORMATION:
; APPLICANT: Efferitz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
; PRIOR FILING DATE: 2001-01-05
; PRIOR APPLICATION NUMBER: US 09/760,427
; PRIOR FILING DATE: 2001-01-13
; PRIOR APPLICATION NUMBER: US 09/855,768
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 4593
; SEQ ID NO 472
; LENGTH: 264
; TYPE: DNA
; ORGANISM: Glycine max
US-09-969-373-472

```

Alignment Scores:

```

Pred. No.: 7.83e-06 Length: 264
Score: 123.00 Matches: 26
Percent Similarity: 61.11% Conservative: 18
Best Local Similarity: 36.11% Mismatches: 24
Query Match: 5.17% Indels: 4
DB: 10 Gaps: 2

```

US-09-890-220-2 (1-445) x US-09-969-373-472 (1-264)

```

QY      370 LeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleuGlu 389
DB      259 ATTGGAGTCT-----GTCCCAATTGACTATAAATAAAAGTTTAATATCTC--- 209
QY      390 AsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerValAspArgPro 409
DB      208 ---TGTTTTAATATATATATATATATATATATATATATATATATATATATATAT 152
QY      410 SerAspSerAsnThrAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLys 429
DB      151 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 92
QY      430 AsnAsnValAspAsnLysAspAsnAsnSerArgAsp 441
DB      91 AATAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 56

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:44:28 ; Search time 3000 seconds
(without alignments)
3729.469 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONCAKSSPEEYISTDE.....INNNKNNVDKNNKNSRDVKIVK 445

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 24791104 seqs, 12571243825 residues
Total number of hits satisfying chosen parameters: 49582208

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=tlh
-Q=/cgn2_1/USPTO.spool/US09890220/runat_13062003_144853_16527/app.query.fasta_1.583
-DB=Pending_Patents_NA_Main -OFMT=fastap -SUFFIX=trpm -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -ONITS=bits -START=1 -END=1 -MATRIX=bloms62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pcr -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09890220.ecgn_1.1.2950.etrnat_13063003_144853_16527 -NCPU=6 -ICPU=3
-NO_MAP -LARGEOUTER -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: Pending Patents_NA_Main:*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq:*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq:*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq:*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq:*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq:*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq:*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq:*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq:*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq:*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq:*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq:*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq:*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq:*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq:*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq:*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq:*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq:*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq:*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq:*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq:*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq:*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq:*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq:*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq:*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq:*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq:*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq:*

29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq:*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq:*
31: /cgn2_6/ptodata/1/pna/US098A_COMB.seq:*
32: /cgn2_6/ptodata/1/pna/US098B_COMB.seq:*
33: /cgn2_6/ptodata/1/pna/US098C_COMB.seq:*
34: /cgn2_6/ptodata/1/pna/US099A_COMB.seq:*
35: /cgn2_6/ptodata/1/pna/US099B_COMB.seq:*
36: /cgn2_6/ptodata/1/pna/US099C_COMB.seq:*
37: /cgn2_6/ptodata/1/pna/US099D_COMB.seq:*
38: /cgn2_6/ptodata/1/pna/US100A_COMB.seq:*
39: /cgn2_6/ptodata/1/pna/US100B_COMB.seq:*
40: /cgn2_6/ptodata/1/pna/US101A_COMB.seq:*
41: /cgn2_6/ptodata/1/pna/US101B_COMB.seq:*
42: /cgn2_6/ptodata/1/pna/US102A_COMB.seq:*
43: /cgn2_6/ptodata/1/pna/US102B_COMB.seq:*
44: /cgn2_6/ptodata/1/pna/US6000_COMB.seq:*
45: /cgn2_6/ptodata/1/pna/US6001_COMB.seq:*
46: /cgn2_6/ptodata/1/pna/US6002_COMB.seq:*
47: /cgn2_6/ptodata/1/pna/US6003_COMB.seq:*
48: /cgn2_6/ptodata/1/pna/US6004_COMB.seq:*
49: /cgn2_6/ptodata/1/pna/US6005_COMB.seq:*
50: /cgn2_6/ptodata/1/pna/US6006_COMB.seq:*
51: /cgn2_6/ptodata/1/pna/US6007_COMB.seq:*
52: /cgn2_6/ptodata/1/pna/US6008_COMB.seq:*
53: /cgn2_6/ptodata/1/pna/US6009_COMB.seq:*
54: /cgn2_6/ptodata/1/pna/US6010_COMB.seq:*
55: /cgn2_6/ptodata/1/pna/US6011_COMB.seq:*
56: /cgn2_6/ptodata/1/pna/US6012_COMB.seq:*
57: /cgn2_6/ptodata/1/pna/US6013_COMB.seq:*
58: /cgn2_6/ptodata/1/pna/US6014_COMB.seq:*
59: /cgn2_6/ptodata/1/pna/US6015_COMB.seq:*
60: /cgn2_6/ptodata/1/pna/US6016_COMB.seq:*
61: /cgn2_6/ptodata/1/pna/US6017_COMB.seq:*
62: /cgn2_6/ptodata/1/pna/US6018_COMB.seq:*
63: /cgn2_6/ptodata/1/pna/US6019_COMB.seq:*
64: /cgn2_6/ptodata/1/pna/US6020_COMB.seq:*
65: /cgn2_6/ptodata/1/pna/US6021_COMB.seq:*
66: /cgn2_6/ptodata/1/pna/US6022_COMB.seq:*
67: /cgn2_6/ptodata/1/pna/US6023_COMB.seq:*
68: /cgn2_6/ptodata/1/pna/US6024_COMB.seq:*
69: /cgn2_6/ptodata/1/pna/US6025_COMB.seq:*
70: /cgn2_6/ptodata/1/pna/US6026_COMB.seq:*
71: /cgn2_6/ptodata/1/pna/US6027_COMB.seq:*
72: /cgn2_6/ptodata/1/pna/US6028_COMB.seq:*
73: /cgn2_6/ptodata/1/pna/US6029_COMB.seq:*
74: /cgn2_6/ptodata/1/pna/US6030_COMB.seq:*
75: /cgn2_6/ptodata/1/pna/US6031_COMB.seq:*
76: /cgn2_6/ptodata/1/pna/US6032_COMB.seq:*
77: /cgn2_6/ptodata/1/pna/US6033_COMB.seq:*
78: /cgn2_6/ptodata/1/pna/US6034_COMB.seq:*
79: /cgn2_6/ptodata/1/pna/US6035_COMB.seq:*
80: /cgn2_6/ptodata/1/pna/US6036_COMB.seq:*
81: /cgn2_6/ptodata/1/pna/US6037_COMB.seq:*
82: /cgn2_6/ptodata/1/pna/US6038_COMB.seq:*
83: /cgn2_6/ptodata/1/pna/US6039_COMB.seq:*
84: /cgn2_6/ptodata/1/pna/US6040_COMB.seq:*
85: /cgn2_6/ptodata/1/pna/US6041_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2378	100.0	1722	US-09-890-220-1	Sequence 1, Appl1
2	2378	100.0	1722	US-09-890-475-57	Sequence 57, Appl1
3	2363	99.4	1722	US-09-890-220-47	Sequence 47, Appl1
4	2285	96.1	1715	US-09-890-220-4	Sequence 4, Appl1
5	2285	96.1	1715	US-09-890-475-58	Sequence 58, Appl1
6	2258.5	95.0	1737	US-09-890-220-7	Sequence 7, Appl1

7	1923	80.9	1497	19	US-09-513-996A-69038	Sequence 69038, A
8	1813.5	76.3	6338	33	US-09-890-220-3	Sequence 3, App11
9	1659	69.8	5895	33	US-09-890-220-6	Sequence 6, App11
10	1639	69.8	201471	20	US-09-534-859-587	Sequence 587, App
11	1639	69.8	201471	31	US-09-803-736-587	Sequence 587, App
12	1166	49.0	1722	42	US-10-219-999-5373	Sequence 5373, App
13	923	38.8	2280	17	US-09-339-947A-2	Sequence 2, App11
14	907.5	38.2	2346	17	US-09-303-460-9	Sequence 9, App11
15	906.5	38.1	2248	17	US-09-339-947A-9	Sequence 9, App11
16	888.5	37.4	2249	74	US-09-303-460-1	Sequence 5, App11
17	871	36.6	2249	74	US-09-303-460-5	Sequence 5, App11
18	868.5	36.5	2226	74	US-09-303-460-7	Sequence 7, App11
19	833	35.0	1841	42	US-10-219-999-16339	Sequence 16339, A
20	833	35.0	1841	75	US-09-312-544-4687	Sequence 4687, App
21	833	35.0	1856	33	US-09-865-439A-20106	Sequence 20106, A
22	833	35.0	2145	74	US-09-303-460-13	Sequence 13, App1
23	782.5	32.9	2296	74	US-09-303-460-11	Sequence 11, App1
24	654	27.5	2640	17	US-09-398-237-6	Sequence 6, App11
25	629	26.5	677	25	US-09-654-617-26623	Sequence 26623, A
26	629	26.5	677	27	US-09-684-016-26623	Sequence 3, App11
27	596.5	25.1	1176	74	US-09-303-460-3	Sequence 3, App11
28	591.5	24.9	990	25	US-09-654-617-386601	Sequence 386601, A
29	591.5	24.9	990	27	US-09-684-016-386601	Sequence 17, App1
30	559	23.5	704	74	US-09-303-460-17	Sequence 17, App1
31	527	22.2	600	74	US-10-260-238-2870	Sequence 2870, App
32	527	22.2	600	61	US-09-325-448-2870	Sequence 179, App
33	527	22.2	600	76	US-09-325-448-2870	Sequence 2870, App
34	521	21.9	492	43	US-10-260-238-1386	Sequence 1386, App
35	521	21.9	492	76	US-09-325-448-1386	Sequence 1386, App
36	510	21.4	726	43	US-10-260-238-2869	Sequence 2869, App
37	510	21.4	726	76	US-09-325-448-2869	Sequence 2869, App
38	492	20.7	2439	42	US-10-219-999-29250	Sequence 29250, A
39	489	20.6	531	33	US-09-874-708A-29253	Sequence 29253, A
40	489	20.6	531	65	US-09-211-750-19028	Sequence 29028, A
41	472.5	19.9	443	63	US-09-167-872-14653	Sequence 14653, A
42	455	19.1	558	61	US-09-170-912-180	Sequence 180, App
43	451	19.0	453	33	US-09-890-220-13	Sequence 13, App1
44	451	19.0	453	61	US-09-172-375-387	Sequence 387, App
45	443	18.6	566	33	US-09-865-439A-88732	Sequence 88732, A

ALIGNMENTS

RESULT 1
US-09-890-220-1
Sequence 1, Application US/09890220
GENERAL INFORMATION:
APPLICANT: Dean, Caroline
TITLE OF INVENTION: Methods and means for modification of plant characteristics using
FILE REFERENCE: Mewburn
CURRENT APPLICATION NUMBER: US/09/890,220
PRIOR FILING DATE: 2001-07-27
CURRENT APPLICATION NUMBER: PCT/GB00/00248
PRIOR FILING DATE: 2000-01-28
PRIOR APPLICATION NUMBER: GB 9901927.5
NUMBER OF SEQ ID NOS: 77
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 1722
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-890-220-1

Alignment Scores:
Fred. No.: 3.72e-230
Score: 2378.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 33
Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-220-1 (1-1722)	
QY	1 MetCysArgGlnAsnGlySerProGluValIleSerThrAspGlu 20
DB	231 ATGTGTAGCGAGAAATTCGCCGCAAAATCCACCGAGAAATGATTTCAACTATGAG 290
QY	21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB	291 AATCTCTGAATATTTGTAATGTAACCTTCGACATATATATACCTTCACTGCTCTCA 350
QY	41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB	351 GGCACCCATGCTTCTTCACNAGATGCTTCACTACAAATTTGACCAAGCGCAAGAA 410
QY	61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnSerThrLeuGln 80
DB	411 AAGTCAAGATCTACGGATGATGATTTCACTATAGGATTTGATTAACACTTACAG 470
QY	81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB	471 AAAATGAAAGTAAAGGAGAGATTTCTTGTGCAATTTGCTCATTTGCTATGCTGATGCTTC 530
QY	101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB	531 AAGGGCGTCGAATTTTCATTTCAATTCATTCATGATTTTCAATTTGATTCAGTCAAGCTT 590
QY	121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB	591 TTCGAGAAATCCACACAGCTTAATGTTCTGTAACCTTCAATTTCAATTTGAGGAA 650
QY	141 GluGlySerAspAspAspLysPheGluProPheSerLeuGlySerLysProArgLysArg 160
DB	651 GAGGAGAAATGATGAGAAATTTGACCCCTTCTCTGCTGCGCAAACTGCTTAAGCGG 710
QY	161 ArgGlnArgGlyGlyArgAsnSerThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB	711 AGACAAAGATGTCAGAAATTAACCCAGGAGACTTAAGTAACTTTTACCGTTGAT 770
QY	181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArg 200
DB	771 TCACCCAGCTTAATTAATGCGCAGAAATGAGATCACTTATATATGAAACCGT 830
QY	201 GlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
DB	831 GGTTTAGGATATCCCGAGCGCAACAGACTTGTGAGCAATTTGATGACCGCAACATTT 890
QY	221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB	891 CCACGAGCAATAGCCCACTCTCTGAGCGTGTGCTAAAGTTATTTGACCAAGCGAA 950
QY	241 AlaValValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB	951 GGTGTGTCCTGCTACTAAGACAAAGAAATTTATCTGTGAGGATCAGAGCTTGAAGC 1010
QY	261 HisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB	1011 CACCTACTTCTTCAGAAAGCGCAATTTATCATCTTCACAGATCCAGCAATGGCGCTT 1070
QY	281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
DB	1071 GAGCAGAGTAATGTGTGACCGGATAGGAGATGAAGATGATGATGATGATGATGATTTT 1130
QY	301 GluAspArgGlnMetLeuAspAspPheValAlaAspValAsnLysAspGluLysGlnPheMet 320
DB	1131 GAAAGATCCCGAGATGCTTGTATGACTTTTGTGAGTGAATTAAGATGAATTAATTCATG 1190
QY	321 HisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspCylHisIleSerThr 340
DB	1191 CATCTTTCGAACTGCTTGTAGCAAAACAAAGGCTTATAGCAGATGCTATATCTCTGG 1250
QY	341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360

Db 1251 GCATGTGAACATTTTCAAGATTCTTACGAAAGAGTTCACCGTTCATCATCTCTTC 1310
 QY 361 TTPCYSTRPARGLEUHEULEULEYLSLEUTPRASNIGLYLEUVALSPSERIATPR 380
 Db 1311 TGGTGTGAGATGTTTTGTGATTAACATGACACGAGCTTGCTCAGTCAGCCACC 1370
 QY 381 ILEASNASCYSASNTHIRILEUGLUAANCYSARGASERSEASERSPHTRHTRHTR 400
 Db 1371 ATCAACAATGCAATTCATCTCCGAGAAATGCCGTAATAGCTCAGACACACACCCACC 1430
 QY 401 ASNASNASNSENSEVALSPARGPROSERASPSERASNTHRASNASNASNILEVAL 420
 Db 1431 AACCAACAACACAGTGTGATGCTCCAGTACTCAACACCAACAATTAACATTTG 1490
 QY 421 ASPHISPROASNAPRILEASNASNLYSASNASNVALASPSNLYSPASNSENSEARG 440
 Db 1491 GATCATCCCAATGACATTAACAACAAGAAATGTTGACACAAAGACAAATTAACACGAGA 1550
 QY 441 ASPLYSVALILELYS 445
 Db 1551 GACAAAGATTAATTA 1565

RESULT 2

US-09-890-475-57
 / Sequence 57, Application US/09890475
 / GENERAL INFORMATION:
 / APPLICANT: Johanson, Urban
 / APPLICANT: West, Joanne
 / APPLICANT: Dean, Caroline
 / TITLE OF INVENTION: Arabidopsis thaliana derived FrigIda gene conferring late flowering
 / FILE REFERENCE: Mewburn
 / CURRENT APPLICATION NUMBER: us/09/890,475
 / CURRENT FILING DATE: 2001-11-13
 / PRIOR APPLICATION NUMBER: PCT/GB00/00197
 / PRIOR FILING DATE: 2000-01-25
 / PRIOR APPLICATION NUMBER: GB 9902660.1
 / PRIOR FILING DATE: 1999-02-05
 / NUMBER OF SEQ ID NOS: 58
 / SOFTWARE: PatentIn Ver. 2.1
 / SEQ ID NO 57
 / LENGTH: 1722
 / TYPE: DNA
 / ORGANISM: Artificial Sequence
 / FEATURE:
 / OTHER INFORMATION: Description of Artificial Sequence: Landsberg
 / US-09-890-475-57

Alignment Scores:

Pred. No.: 3,72e-230 Length: 1722
 Score: 2378.00 Matches: 445
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 33 Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-475-57 (1-1722)

QY 1 METCYSTRPARGLEUHEULEULEYLSLEUTPRASNIGLYLEUVALSPSERIATPR 20
 Db 231 ATGTGAGCAGAAATGTCGCGGAAATCTCACCCGAGAAAGTGTTCACATGATGAG 290
 QY 21 ASNLEULEULEYCYSLYSPROVALARGLEUTPRASNILEPHEHISTLEUARGSELEU 40
 Db 291 AATCTTTATATATTGTAACCTGTCGACTATATACATCTTTACCTGCTCTCTA 350
 QY 41 GLYASNPROSEPRHELEUPROARGCYSLASNLYRYSILEGLYALALYSARGLYSARG 60
 Db 351 GCGAACCCATCGTTTCTCCAGAGTCTGAAGTCAAAAATTGAGCAAAAGCGCAAAAGA 410
 QY 61 LYSERARGSERTHIRIGLYEVALVALPHEASNTYRLYSPCYASNASNTHIRLEUGIN 80
 Db 411 AAGTCAAGATCTACTGGGATGTGATTTTCACTAATTAAGGATTTGTAATAACACATTAACAG 470

QY 81 LYSTHRGLVALARGLUAAPCYSSERCYSPROPHECYSEMERLEUCYGLYSERPHE 100
 Db 471 AANAAGTGAAGTTAGGAGAGATTTGTTCTTCATTTTGGCTATGCTATGATGTGACTTC 530
 QY 101 LYSGLYLEUGLNPHEHISTLEUASNSENSEHISASPLEUPHEGLUPHEGLUPHELYLEU 120
 Db 531 AAGGGGCTCAATTCATTGATTCATGATCATGATTTATTTGAATTTGATGTTCAAGCTT 590
 QY 121 PHEGLUGLTYRGLINTHRVALASNVALSERVALYLSLEUASNSEPRHEIPEHEGLUGLU 140
 Db 591 TTCGAAGAATTACAGACAGCTTAATGTTTGTGAAAACCTTAATCCCTTCATATTAGAGA 650
 QY 141 GLUGLISERASPSAPRILYSPHEGLUPROPHESERLEUCYSELYSPROARGLYSARG 160
 Db 651 GANAGAGATGATGACATTAATTTGAGCCCTTCTCTCTGCTCGAACCCTGTAAAGCG 710
 QY 161 ARGGLNARGGLYGLYARGASNSENTHARGARGLEUVALCYSPHELEUPROLEUASP 180
 Db 711 AGACAAAGAGGTGGCAGAAATTAACACAGAGAGCTTAAGTATGCTTTTACCGTTGGAT 770
 QY 181 SERPROSERLEUTHRASNGLYTHRGLUANGLYLERHLEULEUASNAPGLYASNARG 200
 Db 771 TCACCCAGTTTAATGATGACAGAAAATGGAATACCCCTTAATATGGAACACCGT 830
 QY 201 GLYLEUGLYTRPROGLUALATHRGLULEUALAGLGLNPHEGLUETHRSEASNILE 220
 Db 831 GGTTHAGATATCCGAGGCAACAGAGCTTGCTGGCAATTTGAGTGGCCAGCAACATTT 890
 QY 221 PROVALIILEALHISSESERLEUASAPALAGLYALALYVALILEUTHRSEGLU 240
 Db 891 CCACCAAGCATAGCCCACTCTCTCGACGGTGGTCAAAAGTTATATGACAAAGCGAA 950
 QY 241 ALAVALVALPROALATHRILYSTRHARGLYLSLEUSERIALGLUARGSERGLUALARGSER 260
 Db 951 GCTGTGTCCTCCGCTACTAAGACAAAGAAATATATCGCAGACGAGCAAGGCTTAAGAC 1010
 QY 261 HISLEULEULEGILNYSARGINPHEUTYRTHISSENHISARGVALGINPROMETALALEU 280
 Db 1011 CACCTACTCTTCAGAAAAGCCCAATCTATCTATCTCACAGATCCAGCAATGGCGCTT 1070
 QY 281 GLUGLINALMETSERASPARARGSPSERGLUASPGILVALASPSAPVALALASPPHE 300
 Db 1071 GAGCAAGTAAATGCTGACCGGAGTACGAGATGAAGTGCATGACATGTCACAGATTTT 1130
 QY 301 GLUASPARGLIMETLEUASPSPHEVALASPVALLSNLYASPGILUYSGLINPHEMET 320
 Db 1131 GAAGATCGCCAGATGCTGTGATGACCTTGTGATGTAATTAAGATGAAGCAATTCATG 1190
 QY 321 HISLEUTPRASNSEPRHEVALARGLYGINARGVALILEALASPGILYHISILESETRP 340
 Db 1191 CATCTTTGAGACTCGTTTGTAGAAAACAAGGTTATAGCAGATGCTATATCTCTTGG 1250
 QY 341 ALACYSGLUALAPHESEARGPHERYRGLYLSGLILEUHNISARGTYRSEPRLEUPHE 360
 Db 1251 GCATGTGAAGCATTTTCAAGATTTTACGAGAAAAGATGTCACCGTTCATCATCTCTTC 1310
 QY 361 TTPCYSTRPARGLEUHEULEULEYLSLEUTPRASNIGLYLEUVALSPSERIATPR 380
 Db 1311 TGGTGTGAGATGTTTTGTGATTAACATGACACGAGCTTGCTCAGTCAGCCACC 1370
 QY 381 ILEASNASCYSASNTHIRILEUGLUAANCYSARGASERSEASERSPHTRHTRHTR 400
 Db 1371 ATCAACAATGCAATTCATCTCCGAGAAATGCCGTAATAGCTCAGACACACACCCACC 1430
 QY 401 ASNASNASNSENSEVALSPARGPROSERASPSERASNTHRASNASNASNILEVAL 420
 Db 1431 AACCAACAACACAGTGTGATGCTCCAGTACTCAACACCAACAATTAACATTTG 1490
 QY 421 ASPHISPROASNAPRILEASNASNLYSASNASNVALASPSNLYSPASNSENSEARG 440
 Db 1491 GATCATCCCAATGACATTAACAACAAGAAATGTTGACACAAAGACAAATTAACACGAGA 1550

OY 441 Asplyvalilelys 445
Db 1551 GACAAAGTAAATTAA 1565

RESULT 3

US-09-890-220-47
Sequence 47, Application US/09890220

GENERAL INFORMATION:

APPLICANT: Dean, Caroline

APPLICANT: Gendall, Anthony

TITLE OF INVENTION: Methods and means for modification of plant characteristics using

FILE REFERENCE: Mewburn

CURRENT FILING DATE: 2001-07-27

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: GB 9901927.5

PRIOR FILING DATE: 1999-01-28

NUMBER OF SEQ ID NOS: 77

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 47

LENGTH: 1722

ORGANISM: Arabidopsis thaliana

US-09-890-220-47

Alignment Scores:

Pred. No.: 1.24e-228 Length: 1722
Score: 2363.00 Matches: 444
Percent Similarity: 99.78% Conservative: 0
Best Local Similarity: 99.78% Mismatches: 1
Query Match: 99.37% Indels: 0
DB: 33 Gaps: 0

US-09-890-220-2 (1-445) x US-09-890-220-47 (1-1722)

OY 1 Metcysarglnasncysatgalaalyserserproglugluvalileserthrspglu 20
Db 231 ATGTGTAGCGAGATTTCTCCGGAATTCCTACCGGAGGAAAGTATTCACACTATAG 290
OY 21 Asnleuleuiletyrcyslyspovalargleutyranlephelhisleuargserleu 40
Db 291 AATCTGTGATATATTTGTAACCTGTTGCACTATATACATCTTTCACCTTGCTCTTA 350
OY 41 Glysnpserpserpheleuproargcysleuasntyllysileglialalysarglysar 60
Db 351 GGCACACCATCGTTCTCCAGATGCTGAACTACAAAATTGGAGCAAAAGCGCAAAAGA 410
OY 61 Lyssearargserthrlymetvalalpheasntyllysaspccysasnansnthreugln 80
Db 411 AAGTCAGATCTACTGTGGATGTAGTTTCACTATATACATCTTTCACCTTGCTCTTA 470
OY 81 Lysrhrgluvalarggluaspcyssercysprophcyssermetleucysglserphe 100
Db 471 AAAACGTAAGTTAGGGAGATGTGTTGTCATTTGCTCATATGTCATGTCATGCTTC 530
OY 101 Lysglileuglnphehisleuasnserserhisaspheuphagluphaglupheylsleu 120
Db 531 AAGGGCTGCAATTTCTTTGATTCATCTCATGATTTTGAATTTGAGTTCAAGCTT 590
OY 121 Phegluglutyrglnthrvalasnvalservalysleuasnserserpheliephglu 140
Db 591 TTCAGAAATACCGAGAGTTAATGTTCTGTAATACTTAATCTTCAATATTGAGAA 650
OY 141 Glugliseraspaspaspaspaspaspaspaspaspaspaspaspaspaspaspaspasp 160
Db 651 GAAAGGAGTATGACGTAATTTGAGCCCTTCTCTCTGCTCAAAACCTCGTAAGCGG 710
OY 161 Argglinargllyglargasnaasnthrargargleuylvalcyspheluproleuas 180
Db 711 AGACAAAGAGGTGCGAANAATACACAGAGAGACTTAAGATGCTTTTACCGCTGAT 770

OY 181 Serproserleuthrasnlythrgluasnlythlethrleuasnaspsgllyasnarg 200
Db 771 TCACCGAGTTAATGATGCGACAGAAATGATATCACCCTACTTATATGATGAAACGCT 830
OY 201 Gylleuglytyrproglualaethrgluuaslaagluphagluphagluphagluphagl 220
Db 831 GGTTTAGATATCCGAGGCAAGAGCTTCTGCAATTTGATGACGACAACTT 890
OY 221 Proproalialalalhisserleuaspsalaglyalalysvalileuthrserglu 240
Db 891 CCACCGGATAGGCCACTCTTCTGACCTGTGCTTAATGTAATTTGACAAACGGA 950
OY 241 Alavalvalproalalthrlysthrarglylsleuseralagluarserglualaargser 260
Db 951 GCTGTGTCCTCTGCTACTAGACAGAAAGTATCTGATGAGATCAGAGGCTAGAACG 1010
OY 261 Hisleuleuleuglnlyslarglnphethrhisserhisargvalglnprometalaleu 280
Db 1011 CACCTACTCTTCAGAAAGCGCAATTCATCTCAGAGATCCAGCCAAATGCGGCTT 1070
OY 281 Gluglnvalmetseraspargaspsercyluaspsgluvalaspaspaspvalalasphe 300
Db 1071 GAGCAAGTAAATGCTGACCGGATAGCAGGATGAAGTCATGACATGTCGACATTTT 1130
OY 301 Gluasparglmetleuaspsasphevalaspvalasnlysaspsglulysglnphemet 320
Db 1131 GAAGATCGCCACATGCTTGATGATCTTGTGATGTAATTAAGTGAAGCAATTCATG 1190
OY 321 Hisleutrpasnsersphevalarglyslinarvalilalaspgllyhisleutrp 340
Db 1191 CATCTTGAATCTGCTTTGTAAGAAACAAAGGCTTATACATATGCTATCTCTTGG 1250
OY 341 Alacysglualaaphserargpheryglulysgluleuhsargtyrserleupe 360
Db 1251 GCATGTGAGCATTTTTCAGATTTTACAGAAAGTGTGACCGCTTTCATCATCTTTC 1310
OY 361 TTPCysTTPARGleupheleuilelysleutrpasnhisglyleuvalaspseralathr 380
Db 1311 TGTGTGTGAGATTTTCTTATTAATCATAGAAACCATGACCTGTGACATCAGCAC 1370
OY 381 Ileasnascysasnthrileleugluasnscysargasnseraspthrthrthrthr 400
Db 1371 ATCAACAACTGCAATACCATCTCGAGATTTGCGTAATGTGCAACACACCCACAC 1430
OY 401 AsnAsnAsnSerValaspargproseraspserasnthrasnasnasnileval 420
Db 1431 AACCAACAAACAGGTGATGATGTCACAGTACTCAACCAACCAATACATATGTCG 1490
OY 421 Asphlaspcoasnaspilasnasnlyasnasnvalaspasnlysaspsasnseratg 440
Db 1491 GATCATCCCAATGACATTAACACAGAACATGTTGACAAAGGACATTAACACAG 1550
OY 441 Asplyvalilelys 445
Db 1551 GACAAAGTAAATTAA 1565

RESULT 4

US-09-890-220-47
Sequence 47, Application US/09890220

GENERAL INFORMATION:

APPLICANT: Dean, Caroline

APPLICANT: Gendall, Anthony

TITLE OF INVENTION: Methods and means for modification of plant characteristics us

FILE REFERENCE: Mewburn

CURRENT FILING DATE: 2001-07-27

PRIOR FILING DATE: 2000-01-28

PRIOR APPLICATION NUMBER: GB 9901927.5

PRIOR FILING DATE: 1999-01-28

NUMBER OF SEQ ID NOS: 77

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 4
 : LENGTH: 1715
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 : US-09-890-220-4

Alignment Scores:

Pred. No.:	1,01e-220	Length:	1715
Score:	2285.00	Matches:	428
Percent Similarity:	97.51%	Conservative:	3
Best Local Similarity:	96.83%	Mismatches:	9
Query Match:	96.09%	Indels:	2
DB:	33	Gaps:	1

US-09-890-220-2 (1-445) x US-09-890-220-4 (1-1715)

```

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB      223 ATGTGTAGGCAAAATTTGCGCGCAAAATCCTCACGAGGAAAGTATTCACATGATGAG 292
QY      21 AsnLeuLeuIleTyrCysAlaProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB      223 ATGTGTAGGCAAAATTTGCGCGCAAAATCCTCACGAGGAAAGTATTCACATGATGAG 292
QY      23 AATCTCTTGATATATTGTAACCTTGTGACTATATTAACATCTTTCACCTTGCTCTCTTA 352
DB      23 AATCTCTTGATATATTGTAACCTTGTGACTATATTAACATCTTTCACCTTGCTCTCTTA 352
QY      41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB      353 GGCACACCATCGTTCTTCGCAAGATGCTTGACACTACAAATTTGGGGCAAAAGCGCAAAAGA 412
QY      61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB      413 AAGTCAAAGATCTACTGCGGATGCTGACTTTTCAACTTAAAGGATTTGATTAATACATTACNA 472
QY      81 LysThrGluValArgGluAspCysSerCysPropheCysSerMetLeuGlySerPhe 100
DB      473 AGAAGTGAAGTATGAGGAGGATGCTTCTGCTTCATTTTGGCTCTATGCTATGCTGAGCTTC 532
QY      101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB      533 AAGGGGCTGCATTTTCATTGAAATTCATCCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592
QY      121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB      593 TTGGAAGATACCAACAGATTAATGTTTCTGTAAACATTAATCTCTCATATATTAGAGNA 652
QY      141 GluGlySerAspAspLysPheGluPheSerLeuPheSerLeuCysSerLysProArgLysArg 160
DB      653 GAAGGAAGTGAATGATGATTAATTTGAGCCCTTCTCTCTCTGCTCGAAGCTGTAAGCT 712
QY      161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
DB      713 AGACAAAGAGGTGGGAGAAATTAACACAGAGACATTAAGTATGCTTTTACCGTTGAT 772
QY      181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyLysArg 200
DB      773 TCACCCAGTTTAACTAATGCGACAGAAATGGAATTTGCCCTGATGATGTAAGAAACCGT 832
QY      201 GlnLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluIleMetThrSerAsnIle 220
DB      833 GGTTTAGGATATCCCGAGGACACAGAGCTGCTGACACAAATTTGAGATGATGCTGACACAT 892
QY      221 ProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
DB      893 CCACCAAGCATAGCCCACTCTCTGACGCTGGGTGAAGTTATATATTAACAACCGAA 952
QY      241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
DB      953 GCTGTGCTCCCTGCTACTAAGCAAGAAAGTTATCTGCTGACGATCAAGGCTGAGAAC 1012
QY      261 HisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAlaLeu 280
DB      1013 CACCTACTTCTTCAAGAAACGCAATTTCTATCTATCTCACAGAGTCCAGCAATGGCGCTT 1072
QY      281 GluGlnValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPhe 300
  
```

```

DB      1073 GAGCAAGTAATGTCGATCGGATAGCGAGATGAAGTGAAGTGAAGATGTCAGATTTT 1132
QY      301 GluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMet 320
DB      1133 GAAGATCGCCAGATGCTTGATGATCTTGTGGATGATGAATTAAGATGAAGCAAAATTCATG 1192
QY      321 HisLeuTyrPheSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
DB      1193 CATCTTTGAACTCGTTTGTATGAAACAAAGGTTATATGACAGATGCTATATCTCTTGG 1252
QY      341 AlaCysGluAlaPheSerArgPheTyrGluLysGluLeuHisArgTyrSerSerLeuPhe 360
DB      1253 GCATGGAAGTATTTTCAAGATTTTACGGAAGAGTGTGACCTGTACTCATCTCTTC 1312
QY      361 TrpCysTrpArgLeuPheLeuIleLysLeuTyrPheAsnHisGlyLeuValAspSerAlaThr 380
DB      1313 TGGTGTGAGATTTGTTTGTATTAACATATGAAACATGAGACTGTGACTCAGCCACC 1372
QY      381 IleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 400
DB      1373 ATCAACACTGCAATATCCATCTCGAAGATTTCCGTAATACCTCA-----GTCACTAAC 1426
QY      401 AsnAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsnIleVal 420
DB      1427 AACCAACACACAGATGTGATCATCCAGTCACTCAACACCAACACAAATTAATTTG 1486
QY      421 AspHisProAspAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsnSerArg 440
DB      1487 GATCATCCGAAATGACATATAAAACAAAGAAATGTTGACMACAGACAAATTAACAGCAGA 1546
QY      441 AspLys 442
DB      1547 GACAAAG 1552

RESULT 5
US-09-890-475-58
: Sequence 58, Application US/09890475
: GENERAL INFORMATION:
: APPLICANT: Johanson, Urban
: APPLICANT: West, Joanne
: TITLE OF INVENTION: Arabidopsis thaliana derived Frigida gene conferring late flow
: FILE REFERENCE: Newburn
: CURRENT APPLICATION NUMBER: US/09/890,475
: PRIOR FILING DATE: 2000-01-25
: PRIOR APPLICATION NUMBER: PCT/GB00/00197
: PRIOR FILING DATE: 1999-02-05
: NUMBER OF SEQ ID NOS: 58
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 58
: LENGTH: 1715
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence: Columbia VRN2
: OTHER INFORMATION: cDNA
US-09-890-475-58

Alignment Scores:
Pred. No.: 1,01e-220 Length: 1715
Score: 2285.00 Matches: 428
Percent Similarity: 97.51% Conservative: 3
Best Local Similarity: 96.83% Mismatches: 9
Query Match: 96.09% Indels: 2
DB: 33 Gaps: 1

US-09-890-220-2 (1-445) x US-09-890-475-58 (1-1715)

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
  
```

Db 233 ATGTGTAGCAGATTGTCGCGGAATCTCACCAGAGAGATGATTTCAACTGATGAG 292
 Oy 21 AsnleuLeuileuYrCysLysProValArgLeuYrAsnIlePheHisLeuArgSerLeu 40
 Db 293 AATCTGTGATATATGTGTAACCTGTTGCATATATTAACATCTTTCACCTCGCTCTTA 352
 Oy 41 GLyAsnProSerPheLeuProArgCysLeuAsnYrLysIleGlyAlaLysArgLysArg 60
 Db 353 GGCACACCATCGTTTCCCAAGATGCTTGAACTACAAATTTGGGGCAAAAGCAAAAGA 412
 Oy 61 LysSerArgSerThrGlyMetValValPheAsnYrLysAspCysArgAsnThrLeuGln 80
 Db 413 AAGTCAGATCTACTCGGATGATGATGTTTCCAACTATAGAGATTGTATATATACATTACAA 472
 Oy 81 LysThrGlnValArgLysPcySerCysProPheCysSerMetLeuCysGlySerPhe 100
 Db 473 AGAAGTGAAGTTAGGGAGATGTTGTTCTTCCATTTTGTCTATGCTATGTGGTAGCTTC 532
 Oy 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLysLeu 120
 Db 533 AAGGGCTGCAATTTTCAATTTGAAATTCATCTCATGATTTATTTGAATTTGAGTTCAAGCTT 592
 Oy 121 PheGluGlnYrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGln 140
 Db 593 TTGGAGAAATACCCAGACAGTTAATGTTCTGTAAACCTTAATTCCTCATATTTGAGGAA 652
 Oy 141 GlnGlySerAspAspLysPheGluProPheSerLeuCysSerLysProArgLysArg 160
 Db 653 GAAGGAAGTATGATGATAAATTTGAGCCCTTCTCTCTGCTGCAAACTCGTAAGCGT 712
 Oy 161 ArgGlnArgGlyLysArgAsnAsnThrArgArgLeuValCysPheLeuProLeuAsp 180
 Db 713 AGACAAAGAGGTGGCGAATAATACACAGACAGACTTAAGAATATGCTTTTACCTTGAT 772
 Oy 181 SerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyLysArg 200
 Db 773 TCACCCAGTTTACTTAATGGCACAGAAATGGAATTTGCCCTCTGTAATGATGCAAAACCT 832
 Oy 201 GlyLeuGlyYrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIle 220
 Db 833 GGTTTAGATATCCCAAGCACAGACAGCTTGTGTGACATTTTGATATGATACAGCAACTT 892
 Oy 221 ProProAlaIleAlaHisSerSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGlu 240
 Db 893 CCACAGCCATACCCCACTCTCTGTGGACGCTGTGCTGAAGATTATTAACAACCGAA 952
 Oy 241 AlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSer 260
 Db 953 GCTGTGTCCTGCTACTAGACAAAGAAAGTTATCTGCTGAGCATGACAGGCTAAGAGC 1012
 Oy 261 HisLeuLeuLeuGlnLysArgGlnPheYrHisSerHisArgValGlnProMetAlaLeu 280
 Db 1013 CACCTACTCTCTCAGAAAGCCCAATTCATATCTCAGACAGTCACAGCCCAATGGCCCTT 1072
 Oy 281 GlnGlnValMetSerAspArgAspSerGlnAspGlnValAspAspValAlaAspPhe 300
 Db 1073 GAGCAAGTATATGTCGATCGGATGAGCAGATGAGATGAGATGAGATGAGATTTT 1132
 Oy 301 GlnAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlnLysGlnPheMet 320
 Db 1133 GAGATCGCCAGATGCTTATGATCTTTGGATGTGATTAAGATGAAAGCAATTCATG 1192
 Oy 321 HisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSerTrp 340
 Db 1193 CATCTTGGAACTCGTTTGTATAGAAAGCAAGGGTTATAGCAGATGCTCATATCTCTGG 1252
 Oy 341 AlaCysGlnAlaPheSerArgPheYrGlnLysGlnLeuHisArgYrSerSerLeuPhe 360
 Db 1253 GCATGTGAAGTATTTTCAAGATTTTACAGAAAGAGTTGACCTGTTACTCTATCTCTTC 1312
 Oy 361 TrpCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyLeuValAspSerAlaThr 380
 Db 1313 TGGTGTGGAGATGTTTGTGATTAACATATGGAACCATGAGCTTGTGACTGACAGCCACC 1372

Oy 381 IleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAsnSerSerAspThrThrThr 400
 Db 1373 ATCAACAACTGCAATATCCATCTCGAGATTTGCCGTATATCTCA -----GTCACTAAC 1426
 Oy 401 AsnAsnAsnAsnSerValAspArgProSerAspSerAspThrAsnAsnAsnIleVal 420
 Db 1427 AACACACACACAGTGTGATCATCCAGTACTCAACACACACACATATATATG 1486
 Oy 421 AspHisProAsnAspIleAsnAsnLysAsnValAspAsnLysAspAsnSerArg 440
 Db 1487 GATCATCCGAATGACATATAAACAAGACAAAGATGTTACAAAGACAAATTAACAGCA 1546
 Oy 441 AspLys 442
 Db 1547 GACAAAG 1552
 RESULT 6
 US-09-890-220-7
 ; Sequence 7, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendall, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; TITLE OF INVENTION: Verbalisation gene VRN2.
 ; FILE REFERENCE: Newburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 7
 ; LENGTH: 1737
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-7
 Alignment Scores:
 Pred. No.: 5e-218 Length: 1737
 Score: 2258.50 Matches: 428
 Percent Similarity: 95.78% Conservative: 3
 Best Local Similarity: 95.11% Mismatches: 9
 Query Match: 94.97% Indels: 10
 DB: 33 Gaps: 2
 US-09-890-220-2 (1-445) x US-09-890-220-7 (1-1737)
 Oy 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGlnValIleSerThrAspGlu 20
 Db 223 ATGTGTAGCAGAAATTTGCGCGGAATCTCACCAGAGAGATGATTTCAACTGATGAG 292
 Oy 21 AsnleuLeuileuYrCysLysProValArgLeuYrAsnIlePheHisLeuArgSerLeu 40
 Db 293 AATCTGTGATATATGTGTAACCTGTTGCATATATTAACATCTTTCACCTCGCTCTTA 352
 Oy 41 GLyAsnProSerPheLeuProArgCysLeuAsnYrLysIleGlyAlaLysArgLysArg 60
 Db 353 GGCACACCATCGTTTCCCAAGATGCTTGAACTACAAATTTGGGGCAAAAGCAAAAGA 412
 Oy 61 LysSerArgSerThrGlyMetValValPheAsnYrLysAspCysArgAsnThrLeuGln 80
 Db 413 AAGTCAGATCTACTCGGATGATGATGTTTCCAACTATAGAGATTGTATATATACATTACAA 472
 Oy 81 LysThrGlnValArgLysPcySerCysProPheCysSerMetLeuCysGlySerPhe 100
 Db 473 AGAAGTGAAGTTAGGGAGATGTTGTTCTTCCATTTTGTCTATGCTATGTGGTAGCTTC 532
 Oy 101 Lys-----GlyLeuGlnPheHisLeuAsnSerSerHisAspLe 113
 Db 533 AAGGTGGCAACTATTACAACTGAGGGGCTGCAATTTTCAATTTGAATTCATCTCATGATTT 592

QY 260 -----SerHisLeuLeuGlnLysArgGlnPheTyrHis 271
 Db 913 ATATTAAACCGAAGAAATTTCTTCACCCACCTTCTTCACAAAGCGCAATTCATAT 972
 QY 272 SerHisArgValGlnProMetAlaLeuGlnValMetSerAspArgAspSerGlnAsp 291
 Db 973 TCTCACAGAGTCCAGCCCAATGGCGCTTGAGCAAGTATCTGACCGGATGCCAGAT 1032
 QY 292 GluValAspAspAspValAlaAspPheGluAspArgGlnMetLeuAspPheValAsp 311
 Db 1033 GAAGTCGATGACGATGCTGTCGATGATTTGAAGATCGGACAGATCTTGATGATCTTGAT 1092
 QY 312 ValAsnLysAspGlnLysGlnPheMetHisLeuTyrPheSerPheValArgLysGlnArg 331
 Db 1093 GTGAATTAAGATGAAAGCAATTCATGATCTTGTGAACTCGTTGTGAAGAAACAAAG 1152
 QY 332 ValIleAlaAspGlnHisLeuSerTyrPheAlaCysGluAlaPheSerArgPheTyrGlnLys 351
 Db 1153 GTTATACAGATGGTATCTATTTCTGGGCGATGGAAGCATTTTCAAGATTTAGAGAA 1212
 QY 352 GluLeuHisArgTyrSerSerLeuPheTyrPheTyrPheArgLeuPheLeuLysLeuTyr 371
 Db 1213 GAGTTGCACCGTATCTATCTCTCTGTTGGAGATGTTTGTATTAATGATAGC 1272
 QY 372 AsnHisGlyLeuValAspSerAlaThrIleAsnAspCysAsnThrIleGlnLysAsnGly 391
 Db 1273 AACCATGACTGTGCGACTGACGCCCATCMACTGCMAATCCATCTCGAGAAATGG 1332
 QY 392 ArgAsnSerSerAspThrThrPheAsnAsnAsnSerValAspArgProSerAsp 411
 Db 1333 CGTAAAGCTCAGACACCCACCAACACAAACAGAGTGGATGCTCCAGTAC 1392
 QY 412 SerAsnThrAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsn 431
 Db 1393 TCANACACCAACCAATTAACATGTGAGTCAATCCCAATGACATTAACCAACAAACAAAT 1452
 QY 432 ValAspAsnLysAspAsnAsnSerArgAspLysValIleLys 445
 Db 1453 GTTGACAAACAGGACATTAACGACGACAAAGATTAATTA 1494
 RESULT 8
 US-09-890-220-3
 ; Sequence 3, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics using
 ; FILE REFERENCE: Mewburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; CURRENT FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; NUMBER OF SEQ ID NOS: 77
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 3
 ; LENGTH: 6338
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-890-220-3

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGlnValIleSerThrAspGln 20
 Db 2342 ATGTAGAGCAGCAAAATTTGGCGCAAAATCCACCGGAGAGATGATTTCAACTGATGAG 2401
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 Db 2402 AATCTCTGATATATTTGTAACCTGTCATATATATTAACCTTTCACTTCCTCTCTTA 2461
 QY 41 GlyAsn----- 42
 Db 2462 GCCAAGATGATTTGCCCTCTCTCTCATCATATATAGTCAAGTAATCTTCACTCTCTG 2521
 QY 42 ----- 42
 Db 2522 TGTAGATCACCACATAATAGTTTGTAGTTGCTAAGCTGATATGATGATCATGCGCA 2581
 QY 43 -----ProSerPhe 45
 Db 2582 GATGTGCTCTTTTGTCTCTTAATTTGAACCTGTGTTGTTGGACGACATCGTTT 2641
 QY 46 LeuProAspCysLeuAsnTyrLysIleGlyAlaLysArgLysArgLysSer----- 62
 Db 2642 CTTCACAGATGCTTAACTACAAATTTGAGCAAAAGCGCAAAAGAAAGTA-TGTTTCTT 2700
 QY 62 ----- 62
 Db 2701 CTGGAATGAGCTGCTACAGTGAATATGTTATTTATCTTACTTCTAATATGGAAGCTGATG 2760
 QY 62 ----- 62
 Db 2761 AACTATTTTATCTTGTGAGTGAATATGACATATGAATGAATGTTTCTTCTTGTTCATG 2820
 QY 63 -----ArgSerThrIleVal 68
 Db 2821 GTATTAACCTTACATTTATTAATAATGTTGTTGTTAGTCAAGTCACTCTGAGATGTA 2880
 QY 69 ValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83
 Db 2881 GTTTTCAACTATTAAGATGTAAATTAACATTAACAGAAATGAAAGTTAGTCTTTTCT 2940
 QY 83 ----- 83
 Db 2941 GTTCTCGACAAATTCGATGCTCAATGCTATGTTCTATGATGATTTGTTATTACTA 3000
 QY 84 -----ValArgGluAspCysSerCysProPheCysSerMetLe 96
 Db 3001 TTTTCTGATGATGTCATGAGTGAAGGAGATGTTCTGTCATTTTGTCTATGCT 3060
 QY 96 UCysGlySerPheLys----- 101
 Db 3061 ATGTGATGCTTCAAGGAGGACACTATTAACAAGTGAAGTTCTCCGGGGCCTTTCATA 3120
 QY 101 ----- 101
 Db 3121 TCTAAGCTGGAATGCTACTGCTGTTTCAATGCTGATTAATCTTCACTGTTGTTACT 3180
 QY 101 ----- 101
 Db 3181 ATTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 3240
 QY 102 -----GlyLeu 104
 Db 3241 GAGAAACATGTCACCATGAGCTTCAACCAATGCTGTTGTTGTTGTTGTTGTTGTTG 3300
 QY 104 LnPheHisLeuAsnSerSerHisAspLeuPheGlnPheLysLeu----- 120
 Db 3301 AATTTCAATTTGAATTCATCATGATGATTTAATTTGAGTTCAAGGATGATGTTTGA 3360
 QY 120 ----- 120
 Db 3361 TCGAATCTGATGTTGCTATGCTTATTAATGAGTTAATGATTAATGATTAATGATTA 3420
 QY 121 -----Phe-GluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPhe 136

Db 3421 CTAATGTAGCTTCGGAGAAATACAGACAGTAAATGTTCTGTAATAACTTATTCCTTC 3480
 QY 137 IIEPHEGLN----- 139
 Db 3481 AATATTGAGGTCAGTACTTAACTGGTTAATGGGAATCCAGTACGTGTGAAAT 3540
 QY 140 -----GluGluGlySerAspAspLysPheG1 149
 Db 3541 TTTGTTAATTCATTCCTTATTGTACTAGGAAGAAGAACTGATGAGATAAATTGA 3600
 QY 149 uProPheSerLeu----- 153
 Db 3601 GCCCTTCTCTCTGTGTAACCTCAGAACCCCTTCGATTAATACCTTAATAGCAGTAAC 3660
 QY 154 -----CysSerly 156
 Db 3661 TCCCTGCTCTCTGTGTCAGTACATCTCTGTAATCCAACCAATTAATGTTTGCAGCTCGAA 3720
 QY 156 sProAlGlySarGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPhe 176
 Db 3721 ACCTGTAGAGCGGAGAGCAAGAGGTGGCAAAATAACACAGAGACTTAAGTATGCTT 3780
 QY 176 eLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAs 196
 Db 3781 TTTACCGTTGGATCACCCAGTTTAACTAATGACAGAAATGGAATCACCCTACTTAA 3840
 QY 196 nAsp----- 197
 Db 3841 TCATGTGAATAATCATATCTTCTGTGCGTTCCCTTGCGCTAGAACCTTCATATTACAG 3900
 QY 197 ----- 197
 Db 3901 AAGAGATACATAGGCGCTGATGTTAGTTTGTACTTCCTCGCATCTTCTTGCA 3960
 QY 198 -----GlyAsnArgGlyLe 202
 Db 3961 GGGTATTGTACAGAACTGATGTACAAATTAATGCGATGCTACAGGAACCGGTGTT 4020
 QY 202 uGlyArgProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProPr 222
 Db 4021 AAGATATCCCGAGGCAACAGACGATCTGTGGCAATTTGAGATGACGAGCAATTCACCC 4080
 QY 222 oAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSerGluAlaVa 242
 Db 4081 ACCCAATAGCCCACTCTCTCTGACCGCTGTCTAAAGTTATATTGACAGGGAAGCTGT 4140
 QY 242 lValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg----- 259
 Db 4141 GGTCCCTGCTACTAGACAAAGAAAGTTATCTCTGAGCGATCAGAGGCTAGAAGGTTGT 4200
 QY 259 ----- 259
 Db 4201 TCATCATGACACCCCGTCATCATTAATACCAATTCCTGTTTACAATGTTCTTCCTATT 4260
 QY 260 -----SerHisLeuLe 263
 Db 4261 ATGGATAGCTGTTTATAGTACTGTCATATTAACGAGAAATTTCTTCAGCACCCTACT 4320
 QY 263 uLeuGlnLysArgGlnPheGlyHisSerHisArgVal----- 275
 Db 4321 TCTTGAGAAACGCCCAATTCATATCTTCACAGAGCCAGTGATCCAGATTCTTCACCC 4380
 QY 275 ----- 275
 Db 4381 TACTCTTAGAGCAATTTCTTAAATGCTCATGARGATATCTTATCAAGCATACTTGCT 4440
 QY 275 ----- 275
 Db 4441 TTGTCTCATCAAAATTTGTATTGTGATGCTATGATGATCAAGCAAAATAGTATGCTCA 4500
 QY 276 -----Gln-Prometa 279
 ||| ||| |||

Db 4501 TGTGTCTCCGTTTTTATTTCACATAACCAAAAAATGATGTTTCTGTGACAAGCCAAATG 4560
 QY 279 lAlaGluGlnValMetSerAspArgAspSerGluAspLysValAspAspAspValAla 299
 Db 4561 CGCTTGAGCAATGATATGCTGTGACGGGATAGGAGATGAGTACGATGACAGATGTTGCAG 4620
 QY 299 sPpHeGluAspArg----- 303
 Db 4621 ATTTGAAAGATGCGCAGATTCATGATTTCTTCTGTGCTCATTAATATAGCAACAGA 4680
 QY 304 -----GlnMetLeu 306
 Db 4681 AATGTATATGATGTAACCTGCTAATGCGCTTTTGAACCTTAAAAAAGCTGCAGATGCTT 4740
 QY 307 AspAspPheValAspValAsnLysAspGlyLysGlnPheMetHisLeuThrAsnSerPhe 326
 Db 4741 GATGACTTGTGTGATGTGATTAAGATGAAGAAAGCAATTCATCATCTTTGGAACTGTTT 4800
 QY 327 ValArgLysGlnArgVal----- 332
 Db 4801 GTAGAAACAAAGGTA-ACACTTCTCTTACACATGAGACACACAAAAAGACCTTANG 4859
 QY 332 ----- 332
 Db 4860 TCTTACATTCATACCTGCTGTAATGATTTGCTTATGGAACCTTGACGCTCAATTATGAT 4919
 QY 333 -----lIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
 Db 4920 TGTGTATGTTTACAGGTTTATACAGATGATATTTCTTGGCAGATGATGACATTTTGA 4979
 QY 347 ArgPheTyGlyLysGluLeuHisArgTySerSerLeuPheThr----- 361
 Db 4980 AGATTTTACGAAAGAGTGTGACCGCTTCTCATCTCTTGTGTAAATATAGTACACC 5039
 QY 361 ----- 361
 Db 5040 AACCATATACAGACACATACACTATCAATCTGTTTCGTTTCGAAAAAATA 5099
 QY 362 -----CysThrArgLeuPheLeuIleLysLeuThrAsnHisGlyLeuValAsp 377
 Db 5100 AAAATTTCCAGGTGTGGAGATGTTTGTGATTAACATAGGAGACCATGACTGTGCGAC 5159
 QY 378 SerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThr 397
 Db 5160 TCAGCCACCATCAACACATGCAATCCATCCGAGAAATTTGCCGTATATGCTCAGACACC 5219
 QY 398 ThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsnThrAsnAsnAsn 417
 Db 5220 ACCACCACCAACAAACAAACAGTGTGATGCTCCAGTCAAGCTCAAAACCAACAT 5279
 QY 418 AsnIleValAspHisProAsnAspIleAsnAsnLysAsnAsnValAspAsnLysAspAsn 437
 Db 5280 AACATTTGTGRTATCCCAATGACATTAACAAACAGAAATGTTGACAAAGGACAT 5339
 QY 438 AsnSerArgAspLysValIleLys 445
 Db 5340 AACAGCAGAGCAAGTATTA 5363
 ||| ||| |||

RESULT 9
 US-09-890-220-6
 ; Sequence 6, Application US/09890220
 ; GENERAL INFORMATION:
 ; APPLICANT: Dean, Caroline
 ; APPLICANT: Gendell, Anthony
 ; TITLE OF INVENTION: Methods and means for modification of plant characteristics us
 ; FILE REFERENCE: Newburn
 ; CURRENT APPLICATION NUMBER: US/09/890,220
 ; PRIOR FILING DATE: 2001-07-27
 ; PRIOR APPLICATION NUMBER: PCT/GB00/00248
 ; PRIOR FILING DATE: 2000-01-28
 ; PRIOR APPLICATION NUMBER: GB 9901927.5
 ; PRIOR FILING DATE: 1999-01-28


```

OY 260 -----Se 260
DB 3821 CTTCCTATTATGATAGTATGTTTACTGCTACGTCATATTAACCGAATAATTTCTTCAG 3880
OY 260 rHisLeuLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
DB 3881 CCACCTACTCTTCAAGAAACGCCAATTTATCATCTCTCACAGATCCAGGTGATCCAGT 3940
OY 276 ----- 276
DB 3941 TCCCTTACCCTACTCTTAGGCATTTTCTTAAATTCGTCATGATGATATCTTATCAAGC 4000
OY 276 ----- 276
DB 4001 ATACTGTGTTGTCTCATCTAATTTGTATTTTGTATCTGATGATATCAAGCAAAAA 4060
OY 276 ----- 276
DB 4061 ATTATGTCATGTGTCTCCGTTTATGTCACACTAACCAAAAACTCATGTTTCTTGTGA 4120
OY 277 -----PrometAlaLeuGlnGlnValMetSerAspArgSerGluValAspAs 295
DB 4121 CAAGCCAAATGGCGCTTGACCAAGTAATGTCGATCGGATAGCGAGATGAAGTGCATGA 4180
OY 295 PASPYAlaAspPheGluAspArg----- 303
DB 4181 CGATGTGCGAGATTTTGAAGATCCGACGATTTCCATGATTTCTTTCGCTTCATTAG 4240
OY 303 ----- 303
DB 4241 TAGGCACAGAAATGGTATACGATGTAACCTGCTTAAGCTTTTGAACCTTAAAAAAGC 4300
OY 304 ---GlnMetLeuAspAspPheValAspValAsnLysAspGlnPheMetHisLeuT 323
DB 4301 TCGAGTCTGTATGCTTGCTTGATGATGTAATTAAGTAAAGCAATTCATGCTCTT 4360
OY 323 rPAsnSerPheValArgLysGln----- 330
DB 4361 GGAGCTCGTTTGAAGAAACAAAGTAACCTTCTTACACTTGAACACACACAAAA 4420
OY 330 ----- 330
DB 4421 AGACCTTATGCTTACATTCATACCTGTCCTAATGATTCGTGTAATGAACTTGAAGCT 4480
OY 331 -----ArgValIleAlaAspGlnHisIleSerThrAlaCysGln 343
DB 4481 CAATTTATGATGATGTTTGCAGGGTTAAGCAGATGTCATATCTCTTGGCATGTGA 4540
OY 344 AlaPheSerArgPheTyrGlnLysGlnLeuHisArgTyrSerSerLeuPheTyr 361
DB 4541 GATATTTCAAGATTTTACAGAAAGAGTTCGACTGTACTCATCACTCTTCTGTAATAT 4600
OY 361 ----- 361
DB 4601 AAGTACCCAAACATATACAGACACATACATCAATTTTCTGTTTTCTGA 4660
OY 362 -----CysThrArgLeuPheLeuIleLysLeuThrPAsnHisGln 374
DB 4661 AAGAAAAATTAATAATTTCCAGGGTGTGGAGATTTGTTGATTAACCTATGGAACCATG 4720
OY 374 yLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGlnAsnCysArgAsn 394
DB 4721 ACTGTGCACTGACCGACCATCAACAACGCAATACATCTCTGAAATTTGCCGTATATC 4780
OY 394 rSerAspThrThrThrThrAsnAsnAsnSerValAspArgProSerAspSerAsn 414
DB 4781 CTCAC-----GTCACCTACACAAACAAACAGTGTGATCATCCAGTACCAACAAAC 4834
OY 414 rAsnAsnAsnAsnIleValAspHisProAsnAspIleAsnAsnLysAsnValAspAs 434
DB 4835 CAACACAAATTAACATTTGTGATCATCCGAATGACATTAATAAACAACAAAGATGTGACA 4894
OY 434 nLysAspAsnAsnSerArgAspLys 442

```

```

DB 4895 CAAGACAAATTAACAGCAGACAG 4919
RESULT 10
US-09-534-859-587
; Sequence 587, Application US/09534859
; GENERAL INFORMATION:
; APPLICANT: Bush, David F.
; APPLICANT: Last, Robert L.
; APPLICANT: Levin, Irene M.
; APPLICANT: Norris, Susan R.
; APPLICANT: Parmell, Laurence D.
; APPLICANT: Rounsley, Steven D.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: PLANT POLYMORPHIC MARKERS AND USES THEREOF
; FILE REFERENCE: 38-10(15493)B
; CURRENT APPLICATION NUMBER: US/09/534, 859
; CURRENT FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 1127
; SEQ ID NO 587
; LENGTH: 201471
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-534-859-587

Alignment Scores:
Score: No.: 2,39e-154 Length: 201471
Percent Similarity: 1659.00 Matches: 424
Best Local Similarity: 42.25% Conservative: 4
Query Match: 41.86% Mismatches: 12
Gaps: 69,768 Indels: 577
Gaps: 20 Indels: 14

US-09-890-220-2 (1-445) x US-09-534-859-587 (1-201471)
OY 1 MetCysArgGlnAsnAsnCysArgAlaLysSerSerProGlnGlnValIleSerThrAspGln 20
DB 39570 ATGTGAGGACAGAAATTTGTCGCGGAAATCTCCACCGAGGAAAGTATTTCAACTGATGAG 39629
OY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 39630 AATCTCTGATATATTTGTAACCTGTCGACATATTAACATCTTTCACCTGCTCTCA 39689
OY 41 GlyAsn----- 42
DB 39690 GCGAA-CGTATGATTTGGCTTCTCTCATCATTTTACCTTACATTTTCAATCTTCATCTCC 39748
OY 42 ----- 42
DB 39749 TGTGTAGATCACCCACTAATATGATTTGATTTGCTAAGTGAATATGCTGACTCATGGC 39808
OY 43 -----Pro 43
DB 39809 GAGTGTGCTCTCTTTTGTCTCTAATGTAATTTGAACCTGTTGTTGTTGTCAGCCA 39868
OY 44 SerPheLeuProArgCysLeuAsnTyrLysIleGlnAlaLysArgLysArgLysSer 62
DB 39869 TCGTTTCTGCAAGATGCTTGAACACAAATTTGGGCAAGGCAAAAGAAAGTA-TGC 39927
OY 62 ----- 62
DB 39928 GTTTCCTTGAATGATGTTGCCACAGTGATATGTAATTTATCTTACTTCTAATATGAA 39987
OY 62 ----- 62
DB 39988 GCTGATGAACATATTATCTTGTGTGATGATATGACATATGAATGATGATCTTCTTCTT 40047
OY 63 -----ArgSerThrGly 66
DB 40048 TTTCATCTATACACTTATATTTTACAAAATTTGCTTTTGTCTTATAGTCAAGATCTACTGG 40107
OY 67 MetValValPheAsnTyrLysAspCysAsnAsnThrLeuGlnLysThrGlu----- 83

```

Db 40108 ATGGTACGTTTCAACATAAGATGTTGAATTAATACATTCAAAAGACATGAGGTTAGTCT 40167
 QY 83 ----- 83
 Db 40168 TTTTCGTCTTTCGACAAAATTCGATGCTAATGTTCTCTAGATGATTTGTTAT 40227
 QY 84 ----- -ValArgGluaspCysSerCysProPheCysSe 94
 Db 40228 TTACTATTTTTCTGTATTGTGCAGCAGATTAGGGAGATTTCTTGCCATTTTGCTC 40287
 QY 94 melleucysGlySerPheLys----- 101
 Db 40288 TATGCTATGTGTAGCTTCAAGGTGGCAACTATTACAACTGAGGTTTCTCCGGGGCCT 40347
 QY 101 ----- 101
 Db 40348 TTCAATATCAACAGTGAATGCTACTGCGCTTAAATGCTATTAATCTTACTGTTGG 40407
 QY 101 ----- 101
 Db 40408 TTACATATTTTTGTGTGTGTGTTGCTTCTTCTTTTAACTGCTGAGTGTGC 40467
 QY 102 ----- -G 102
 Db 40468 TTATCTGAGAAACATGTTCCAGTTCAGCTTACAAATCCATTTGCTGTATGACAG 40527
 QY 102 ----- 102
 Db 40528 GGTGCAATTTTCATTGAAATTCATCTCATGATTTATTGAATTTGAGTTCAAGGTATG-T 40586
 QY 121 ----- 121
 Db 40587 GGTTTTATGGAATTTCTGTTTCCCTATGCGCTTACTGATAGAGTTATGTTAAAAAG 40646
 QY 122 ----- GluGluTyrGlnThrValAsnValSerValLysLeu 134
 Db 40647 GGTCTTCTCTATTTGCTGCTTTTGGAAATACACAGACAGTATATGTTCTGTAATAACTTA 40706
 QY 134 snSerPheIlePheGlu----- 139
 Db 40707 ATTCCTTCATATTGAGGTCACTTACTTAAACTGGTTAATGGGAAATCCTATAGCTG 40766
 QY 140 ----- GluGluGlySerAspAsp 146
 Db 40767 GTGAAAATTTGGTTTATATCCATCCCTATTGTACTAGAGAAAGAAAGATGATGAT 40826
 QY 147 LysPheGluProPheSerLeuCys----- 154
 Db 40827 AAATTTGAGCCCTTCTCTCTG-GTAACTCTCAGAACCCCTTGATTAATACCTTAATA 40885
 QY 154 ----- 154
 Db 40886 GCAATACCTCTGCTTTCTGTGCTAGTACTTCTCTAATAATCCAACCAATGTTTTGC 40945
 QY 155 ---SerLysProAlaTyrLysArgGlnArgGlyArgAsnAspThrArgArgLeuLys 173
 Db 40946 AGCTCGAAACCTGTAGAGCTAGACAAAGAGGGGAGAAATTAACACAGAGACTTAA 41005
 QY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyTleThr 193
 Db 41006 GTATGCTTTTACCGTTGGATTCACCCACTTTAGCTAATAGCAGCAAAAATGGAATGCC 41065
 QY 194 LeuLeuAsnAsp----- 197
 Db 41066 CTGCTGATATGTGTAATAATCACATCTTCTGTGTGATTTGTTGCTGTAGAACTTC 41125
 QY 197 ----- 197
 Db 41126 ATTTTACAGAAGAATACAAATGCTCGATTGTTAGTTTGTACTTCTCTCGCAATTC 41185
 QY 198 ----- -GlyAs 199
 Db 41186 TTCTTGTAGGTAATGTTACAGAACATGATGTACAAAATTAATGTCATGCTACAGAAA 41245

QY 199 nArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAs 219
 Db 41246 CCGTGTATGAGATATCCCGAGGCAACAGACTTGTGGCAAAATTTGAATGATGACAG 41305
 QY 219 nIleProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThrSe 239
 Db 41306 CATTCACAGCCATAGCCCATCTTCTCTGAGCGCTGGCTGAAGTATATTAAACAC 41365
 QY 239 rGluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluAlaSerGluAlaAr 259
 Db 41366 CGAAGCTGTGTCCTGCTACTAAGACAAAGAAATTTATCTGTGACGATCAGAGGCTAG 41425
 QY 259 ----- 259
 Db 41426 AAGTTTGTTCATCATGACACCCGTCATCATTAATACATACCTGTGTACAAATGTT 41485
 QY 260 ----- -se 260
 Db 41486 CTTCCTATTATGATAGTATGTTTACTGACTGACATATTAACCGAGAAAATTTCTTCAG 41545
 QY 260 rHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGln----- 276
 Db 41546 CCAGCTACTTCTTGAGAAAGCCCAATTCATTCATTCACAGAGTCCAGGTGATCCAACT 41605
 QY 276 ----- 276
 Db 41606 TCCCTCAGCTACTTCTTAGGCATTTCTTAAATGCTCATGATGATCTTATCAAGC 41665
 QY 276 ----- -se 276
 Db 41666 ATACTGTGTTGTTCTCATCTAAATTTGATTTTGATTCGTATGATACAGCAAAA 41725
 QY 276 ----- 276
 Db 41726 ATTAATGCCAGTGTGTCCTGTTTATTTGCACTAACCAAAACTGCATTTCTTGTA 41785
 QY 277 ---PrometaIleuGlnGluValMetSerAspArgAspSerGluAspGluValAspAs 295
 Db 41786 CAAGCCAATGGCGCTTGAGCAAAATGATGTGATCGGGATAGCGAAGATGAAGTGAAGA 41845
 QY 295 pasPValAlaAspPheGluAspArg----- 303
 Db 41846 CGATGTGCAGATTTTGAAGATCCCGCAGATTCATGCATGATTTCTGCTTCATTAAG 41905
 QY 303 ----- 303
 Db 41906 TAGCCACAAGAAATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 41965
 QY 304 ---GlnMetLeuAspAspPheValAspValAsnLysAspGlyLysGlnPheMetHisLeu 323
 Db 41966 TGCAGATGCTTGATGATCTTGTGATGTGAATTAAGATGAAGAAAGCAATTCATGATCTT 42025
 QY 323 rPAsnSerPheValArgLysGln----- 330
 Db 42026 GGAATGCTTTGTAAAGAAACAAAGGTAATGATGATGATGATGATGATGATGATGATGAT 42085
 QY 330 ----- 330
 Db 42086 AGACTTATATGCTTACATTCATACCTGCTGCTAAATGATTTCTGTTATGAACTTGAGCT 42145
 QY 331 ----- ArgValIleAlaAspGlyHisIleSerThrPalacyGlu 343
 Db 42146 CAATATTAATGATGATGTTTGCAGGGATATAGCAAGATGTCATATCTTGGGCAATGTA 42205
 QY 344 AlaPheSerArgPheTyrGlyLysGluLeuHisArgLysSerSerLeuPheThr----- 361
 Db 42206 GTATTTTCMAATTTTACGAGAAAGAGAGTTCCTACTGTTACATCACTTCTGGAATAT 42265
 QY 361 ----- 361
 Db 42266 AAGTACACCAACATATACAGACACATTAATCACTATCAATTTGTTTGTGTTTCTGA 42325


```

OY 155 ---serlyProArtygysArGaGlnAArglyGlyArAsAsnThrArgArgLeuLys 173
DB 40946 AGCTGAAACCTGCTAGAGCTAGACAGAGGAGGAGAAATACACCGAGGAGCTTAAA 41005
OY 174 ValCysPheLeuProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyLeuThr 193
DB 41006 GTATGCTTTTACCGTTGATCCACAGTTTATGCTAAATGACAGAAAATGCAATTC 41065
OY 194 LeuLeuAsnAsp 197
DB 41066 CTGCTGATGATGATGTAATAATCACATCTCTCTGCTGATTCGTTGTGCTTAAACTTC 41125
OY 197 197
DB 41126 ATTTTACAGAGAAGATACAAATGCTGATTTGTTAGTTTGTACTCTCTCTGCTATTC 41185
OY 198 198
DB 41186 TTCTTGTAGAGGTAATGTATACAGAACTGATGACAAATTAATGGCATGCTACAGGAAA 41245
OY 199 199
DB 41246 CCCTGTTAGATGATATCCGAGCAGACAGAGCTGCTGACAAATTTGAGATGACTACGTA 41305
OY 219 219
DB 41306 CATTCACACAGACATACCCACATCTCTCTGAGGCTGCTAAAGTTATTTATTAACAC 41365
OY 239 239
DB 41366 CGAAGCTGTGCTCCCTGCTACTACAAAGAAAGTTATCTGTGAGCATCAGAGGCTAG 41425
OY 259 259
DB 41426 AAGGTTTGTTCATCAGACACCCGCTCATCAATTAATACCATCTGTTGTACAAATGTT 41485
OY 260 260
DB 41486 CTTCCTATATGATAGATAGTGTACTGCTACTGCAATTAATACCGAGAAAATTTCTTCAG 41545
OY 260 260
DB 41546 CCACCTACTTCTTCAGAAAGCCAAATTCATCTCTACAGAGTCCAGGATGATCAAGT 41605
OY 276 276
DB 41606 TCCTTACACTACTCTTAGGCATTTCTTTAAATTCGTCATGATGATATCTTATCAAGC 41665
OY 276 276
DB 41666 ATACTGTGTTTTCATCACTAAATTTGATTTTGTATCTGTATGATACACGCAAAAAA 41725
OY 276 276
DB 41726 ATTATGTCATGTTGCTCCGTTTATTCGCCACTAACCAAAACATGCATGTTCTTGGA 41785
OY 277 277
DB 41786 CAGGCTATGCGCTTGTAGAGATGATGCTGATGCGGATGACGAGATGAAGTCGATGA 41845
OY 295 295
DB 41846 CGATGTTGAGATTTTGAAGATCCGAGATTCATGATTTCTTTGCGCTATTAAG 41905
OY 303 303
DB 41906 TAGGCAACAGAAAATGATATGATGATGCTTAATGCTTTGAAACTTAATAAAGC 41965
OY 304 304
DB 41966 TGCAGATGCTTGTGATGCTTGTGATGATTAAGATGAAGAAAGCAATTCATGATCTTT 42025
OY 323 323
DB 323 rpsAnserPheValArgLysGln----- 330

```

```

DB 42026 GGAACCTGTTGTAAAGAAAGGTAAGTAAGTACTTCTTACACTTGAAACACAGACAAA 42085
OY 330 330
DB 42086 AGACCTTATGCTTACATTCATCTGCTAAATGATTCGCTTATGGAACCTTGAGCT 42145
OY 331 331
DB 42146 CAAATTAATGATGATGTTTGCAGGGTATATGACAGATGCTATTCCTTGGCAATGTGA 42205
OY 344 344
DB 42206 GTATTTTCAAGATTTTACAGAAAGAGTGTGACATGATCATCACTCTTGTGTAAAT 42265
OY 361 361
DB 42266 AAGTACACCAACATATACAGACATATACATCAATTTGTTGTTTCTGTA 42325
OY 362 362
DB 42326 AAGAAAAATAAAAATTCACAGTGTGAGATTTGTTTATTAATGATGACATG 42385
OY 374 374
DB 42386 ACTTGTGACAGCAGCAGCAATCAACATGCAATTCATCTGCAAAATTCGCTAAG 42445
OY 394 394
DB 42446 CTCA-----GTCACTAACCAACCAACAGTGTGATGATCAATTAACAAAGAAATGTGACAA 42500
OY 414 414
DB 42500 CAACAAACATTAACATTTGATGATCCGAAATGATCAATTAACAAAGAAATGTGACAA 42559
OY 434 434
DB 434 nlyAspAsnAsnSerArgAspLys 442
DB 42560 CAAGGACATTAACAGCAGAGACAG 42584
OY 42560 42560
DB 42584 42584
RESULT 12
US-10-219-999-5373
: Sequence 5373, Application us/10219999
: GENERAL INFORMATION:
: APPLICANT: Cao, Yongwei
: APPLICANT: Edgerton, Michael D
: APPLICANT: Hinkle, Gregory J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Liu, Jingdong
: APPLICANT: Stein, Joshua
: TITLE OF INVENTION: CDNA SEQUENCES AND USES FOR PLANT IMPROVEMENT
: FILE REFERENCE: 38-10/527261C
: CURRENT APPLICATION NUMBER: US/10/219, 999
: PRIOR FILING DATE: 2002-08-15
: PRIOR APPLICATION NUMBER: US 60/324, 109
: PRIOR FILING DATE: 2001-09-21
: PRIOR APPLICATION NUMBER: US 60/312, 544
: NUMBER OF SEQ ID NOS: 63520
: SEQ ID NO 5373
: LENGTH: 1722
: TYPE: DNA
: ORGANISM: Glycine max
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (164)..(1450)
: OTHER INFORMATION:
: US-10-219-999-5373
Alignment Scores:
Pred. No.: 3,37e-107
Score: 1166.00
Percent Similarity: 65.65%
Best Local Similarity: 52.95%
Length: 1722
Matches: 242
Conservative: 58
Mismatch: 111

```

Query Match: 49.03% Indels: 46
DB: 42 Gaps: 9

US-09-890-220-2 (1-445) x US-10-219-999-5373 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 164 ATGTCCCGCAAAATTTCTCCGTACACCATGCCGGGAAGAATAATTCACCTGATG 223

QY 21 AsnLeuLeuIleTyrCysLysProValAlaGluTyrAsnIlePheHisLeuArgSerLeu 40
DB 224 ACCCTTGATTATTATGCAAGCCTGTGAACTTACAAATATCTCAATCCGCGCTCTT 283

QY 41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 284 CAAATTCCTCTTCTTCTTGTAGAGATCTTGCGTTAAATAAGCAACCCCTAAAG 343

QY 61 LysSerArgSerThrGlyMetValValPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 344 AGCTTGAGAGCA---GGAATGTGATTTCAATTAAGGACCGCTACCAACATCTTCGA 400

QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 401 AAGACTGAAGTACCGAAGACTTCTTCTTCTGCTTGTGATGCAAGTGTGAAACTT 460

QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
DB 461 AAGGCTTGCGATTTCATCTTGTTCATCATCATGATCATTCACCTTGAGTTCTGGCT 520

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
DB 521 ACTGAAGATTACCAAGCAAGTGAATGCTCTGTAAATAATGAT---ATAATGAGATCA 574

QY 141 GluGlySerAspAsp-----AspLysPheGluProPheSerLeuGlySerLysPro 157
DB 575 GGAATGTTGCTGATGAGTAATTCACCAATGCGAAACCTTCTTCTGCTTCAACCT 634

QY 158 ArgLysArgArgGlnArgLysGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
DB 635 CGAAACGTAGAAGAAAGACCTGTTCAAATGAAACGCCCAATGTAATAATTCCTG 694

QY 178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyTleThrLeuLeuAsnAsp 197
DB 695 GAGTGGACTCACCA-----GAAAGCATACATAATGATTTCTACAAAAAAGATGAT 745

QY 197 ----- 197

DB 746 GATATCTATCTCGCAAGAGAAATGTCTAGAACATCTCTGATGAGAGATTTTC 805

QY 198 -----GlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeu 210
DB 806 CCTAGTGAAGAAATGACGAGAGAAATTTGGTCCGATCATCTCGCACACGAGCAAC 865

QY 211 AlaGlyGlnPheGluMetThrSerAsnIleProPro-----AlaIleAlaHisSer 227
DB 866 CTGAGACATGTAAGTCAAGTTTCCAGATTCAGAGTTCATTCATCCATGCCCAATCT 925

QY 228 SerLeuAspAlaGly---AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246
DB 926 TCTGTGACCTCGAATGTAGTAATCAATGTAAAGTAAATCTCTGCTCTCTCT--- 982

QY 247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuGlnLys 266
DB 983 AAAACAAAGAACTAAGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1042

QY 267 ArgGlnPheTyrHisSerHisArgValGlnPrometAlaLeuGlnValMetSerAsp 286
DB 1043 AGACTTCTTTCACCTACACAGAGTCCAGCCTATGAGCATGAGAACAGTATATACAG 1102

QY 287 ArgAspSerGluAspGluValAspAspArgValAlaAspPheGluAspArgGlnMetLeu 306
DB 1103 CGGTATGATGAAGACGAAGTGTATGACGACATTCGAGATCTTGAAATGAAAGATGCT 1162

QY 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPhe 326
DB 1163 GACGATTTTGTGGATCTTCCAAAGATGAAAAACACCTCATCTCGAATCTCTTT 1222

QY 327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
DB 1223 ATGAGGACCAAAAGGCTGCTAGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1282

QY 347 ArgPheTyrGluLysLeuLeuHisArgTyrSerSerLeuPheThrProCysThrArgLeu 366
DB 1283 AAGCTTCATGAGAAAGAGCTGATCTCATCTTCACGCTTATTTGGTGGAGGTTATTC 1342

QY 367 LeuIleLysLeuThrPheAsnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
DB 1343 ATGATCAAACTTGGATATCATGATGCTCTTGTGATCCCTGTACATGACAACTGTACATA 1402

QY 387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrThrAsnAsnAsnSerVal 406
DB 1403 GTATTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1447

QY 407 AspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisPro 423
DB 1448 TGAAGACCATTAAGAACTGACTCAACAAACAGATATACAGTCTTCTCG 1498

RESULT 13
US-09-339-947A-2
; Sequence 2, Application US/09339947A
; GENERAL INFORMATION:
; APPLICANT: YOSHIDA, Nobumasa
; APPLICANT: YAMAI, Yukihito
; APPLICANT: KATO, Yoshihiro
; APPLICANT: HIRATSUKA, Junzo
; APPLICANT: TAKAHASHI, Shigeru
; APPLICANT: MIWA, Tatsushi
; TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
; FILE REFERENCE: 032735-001
; CURRENT APPLICATION NUMBER: US/09/339,947A
; CURRENT FILING DATE: 1999-06-25
; PRIOR APPLICATION NUMBER: JP 1998-180065
; PRIOR FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: JP 1999-179043
; PRIOR FILING DATE: 1999-06-24
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (310)..(2142)
; NAME/KEY: misc_feature
; LOCATION: (1650)..(1655)
; OTHER INFORMATION: BamHI recognition site
; NAME/KEY: misc_feature
; LOCATION: (1984)..(1989)
; OTHER INFORMATION: SphI recognition site
US-09-339-947A-2

Alignment Scores:
Pred. No.: 2,23e-82 Length: 2280
Score: 923.00 Matches: 216
Percent Similarity: 43.23% Conservative: 52
Best Local Similarity: 34.84% Mismatches: 108
Query Match: 38.81% Indels: 244
DB: 17 Gaps: 9

US-09-890-220-2 (1-445) x US-09-339-947A-2 (1-2280)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
DB 310 ATGTCCCGCAAAAGACTCCGCTGCTGATTTCTGAGAGAGAGAGATTCCTGCTGAAAG 369

TITLE OF INVENTION: GENE FOR FLORAL REGULATION AND METHODS FOR CONTROLLING
 FILE REFERENCE: 032735-001
 CURRENT FILING DATE: 1999-06-25
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: JP 1998-180065
 PRIOR FILING DATE: 1999-06-24
 NUMBER OF SEQ ID NOS: 20
 SOFTWARE: Patentlin Ver. 2.0
 SEQ ID NO 9
 LENGTH: 2248
 TYPE: DNA
 ORGANISM: Oryza sativa
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (86)..(1897)
 NAME/KEY: misc_feature
 LOCATION: (36)..(41)
 OTHER INFORMATION: NheI recognition site
 US-09-339-947A-9

Alignment Scores:
 Pred. No.: 1 03e-80 Length: 2248
 Score: 906.50 Matches: 209
 Percent Similarity: 42.68 Conservative: 59
 Best Local Similarity: 33.28 Mismatches: 109
 Query Match: 38.128 Indels: 251
 Gaps: 7

US-09-890-220-2 (1-445) x US-09-339-947A-9 (1-2248)

QY 1 MetCysArgGlnAsnGlySerProGluGluValIleSerThrAspGlu 20
 DB 86 ATGGCCGCCACACAGGCGCTGCTCTCCCGATGACAGCTTGGCAAGAA 145
 QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 DB 146 AGCTGCATTAATACGCAAGCGGTCGAGTTGTAATATCATTCACGCGCATTCATT 205
 QY 41 GlysAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 DB 206 AAAAATCTCTCTTCTTCTCAAGATGCCCTTCTTCAAGATTCAAGCAAGAGAGAG 265
 QY 61 LysSerArgSerThr----- 65
 DB 266 AGAGCCCTGATTAACCATATACCTTCTGAGGACACAATAAAGAACTCGGGACAAAT 325
 QY 65 ----- 65
 DB 326 ATCTTCCCTTATGTTCTGTAGCTAGACTACTAATAATGTTTCACTTGAAGGCAT 385
 QY 65 ----- 65
 DB 386 TCTCGATATATGATTAGTTCAGTGTGCTGTTGTTGACTCTTTTCATGAAATTTGGAAT 445
 QY 65 ----- 65
 DB 446 AAAGACTACACTGAACCAACATTCGTCATCTCTGATGTGAAGAACTTAGCAACTCCGA 505
 QY 65 ----- 65
 DB 506 GCTTGCAGCCTTAATATTATTCCTTATCAGCTGTGAGAGAGCTGAGCAAACTTTTGATGAC 565
 QY 65 ----- 65
 DB 566 AATACTGTTCTGGGAACCATGTGGAAGGCTCTACTCTCAAAAGCTTGAAGGAAGTGT 625
 QY 65 ----- 65
 DB 626 TTCTGGGTAATAATACCAATCGATCTTCTGCTTCATCTTGGGAAATTTGTGAGCTTA 685

QY 65 ----- 65
 DB 686 AGTTGGACATACCGTGAATATGTCACGCGTGGATGATGACCAAGCTTCTTAGAG 745
 QY 65 ----- 65
 DB 746 CCAAAATTTTCGAGATGACAGTTCCTGACATTTTGTCTCAGAAGGTATGCTACT 805
 QY 65 ----- 65
 DB 806 GGTTCATTTCACTGACAGTATGATCTGCTCAAGAGCGTGTGCAAAAGCATGTCC 865
 QY 65 ----- 65
 DB 866 GAGTCTCCTTATAGTATTATCATATATGATGTCACACTTCGTCATTGACATATT 925
 QY 66 ----- 925
 DB 926 ATAGGTGATGATCGCAATGCTCTTTTACTACAAATACATATATCTATGCA 985
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 986 AAACCGAAGTCACTGAAGATTTCTTCCCATTTTGTGATGATGATGATGATGATGAT 1045
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPhe 120
 DB 1046 AAGGCTTAGAGATGATCACTTAACGATCGCATGACCTTTCCATTTATGATTTGGATA 1105
 QY 121 PheGluGlyTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu 139
 DB 1106 TCTGAAGAGTCCAGCGCTGTTAATGTTAGTCAAGACATCTGTTGGAACAGAGCTT 1165
 QY 140 ---GluGluLysSerAspAspLysPheGluProPheSerLeuGlySerLysProArg 158
 DB 1166 TTGCGTGGAGAGT---GATCCAAAGCATCAAACTTTTCTGACCGCAAGATTTAG 1222
 QY 159 LysArg---ArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
 DB 1223 AAGCGTAAAGAGGGAATCTCAAGTAAATTAAGCATGATCATCATATTGTG 1282
 QY 178 ProLeuAspSerPro---SerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsn 196
 DB 1283 GATTGAGGATCACTGGAAGATGCCAGCGAGTCTGAAGACATTAACCTGAGAGGAA 1342
 QY 197 AspGlyAsnArgGlyLeuGlyTyrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 216
 DB 1343 AATGCTAGT----- 1351
 QY 217 ThrSerAsnIleProProAlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal--- 235
 DB 1352 ----- 235
 QY 236 -----IleLeuThrSerGluAlaValAlaProAlaThrLysThrArgLysSerAla 253
 DB 1394 GGTAGCAATCTTTGACGACCAACAGTTCACATTTGGGAGCAAGAAAGTGTCTGTT 1453
 QY 254 GluArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrHisSerHis 273
 DB 1454 GAACGAGTGAATCCCAAAATACGCAAGCTCTCAAAAGCCAGCTTCTTATTTCTGAC 1513
 QY 274 ArgValGlnProMetAlaLeuGlnGlnValMetSerAspArgAspSerGluAspGluVal 293
 DB 1514 AGGGCTCAACCAAGGATGAGACCAAAAGTTTCTCGAATCGAATGAGATGATGATTT 1573
 QY 294 AspAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsn 313
 DB 1574 GATGATGATCATCTGATTTTGAAGATAGAAAGATCTTGATGATTTGTTGATGTACA 1633
 QY 314 LysAspGluLysGlnPheMetHisLeuThrPheSerPheValArgLysGlnArgValIle 333
 DB 1634 AAAGCGAAGAACTTATATGATGATGGAATTTCTTGTTCGGAACCAAGGATGACTA 1693
 QY 334 AlaAspGlyHisIleSerThrPheLysGluAlaPheSerArgPheTyrGluLysGluLeu 353

```

Db      1694 GCGGATGGCATATTCCTGGGCATGCGAGCATTCTGCACTTCATGACACAGAACTT 1753
QY      354 HISArgTYrSerSerLeuPheTYrPArgLeuPheLeuIleLYsLeuTrpAsnHis 373
Db      1754 GTACAAATCCAGCTCTACTATGAGGTGGAGGTTTTTATGTCAAACTCTGGAACCCAC 1813
QY      374 GlyLeuValAspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
Db      1814 AGTCTACTGGATGCGCGGCATGATGCTGCAACACATCTTGAAGGCTACCTGAAC 1873
QY      394 SerSerAspThrThrThrAsn 401
Db      1874 GGAAGCTCGATCCAAAGAAAAAT 1897

```

Search completed: June 20, 2003, 02:18:49
 Job time : 3231 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 19, 2003, 23:51:48 / Search time 956 Seconds
(without alignments)
3236.070 Million cell updates/sec

Title: US-09-890-220-2
Perfect score: 2378
Sequence: 1 MCRONCRASSPEEIVSYDE.....INNNVNDKNNSHDKYIK 445

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 7816032 seqs, 3476047932 residues
Total number of hits satisfying chosen parameters: 15632064

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO_SPOOL/US09880220/runat_13062003_144853_16556/app.query.fasta.1.583
-DB=ending_Patents_NA_New -OFMT=fastap -SUFFIX=trnp -MINMATCH=0.1 -IOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09880220.ecgn.1.1.844.etrnat.13062003.144853.16556 -NCPU=6 -ICPU=3
-NO_MAP -LARGEREQUY -NEG.SCORES=0 -WAIT -DSPLOCK=100 -LONGLOCK
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/PCIT_NEW.COMB.seq:*
2: /cgn2_6/ptodata/2/pna/US06_NEW.COMB.seq:*
3: /cgn2_6/ptodata/2/pna/US07_NEW.COMB.seq:*
4: /cgn2_6/ptodata/2/pna/US08_NEW.COMB.seq:*
5: /cgn2_6/ptodata/2/pna/US09_NEW.COMB.seq:*
6: /cgn2_6/ptodata/2/pna/US09_NEW.COMB.seq2:*
7: /cgn2_6/ptodata/2/pna/US09_NEW.COMB.seq3:*
8: /cgn2_6/ptodata/2/pna/US10_NEW.COMB.seq:*
9: /cgn2_6/ptodata/2/pna/US10_NEW.COMB.seq2:*
10: /cgn2_6/ptodata/2/pna/US10_NEW.COMB.seq3:*
11: /cgn2_6/ptodata/2/pna/US60_NEW.COMB.seq:*
12: /cgn2_6/ptodata/2/pna/US60_NEW.COMB.seq2:*
13: /cgn2_6/ptodata/2/pna/US60_NEW.COMB.seq3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1166	49.0	1722	US-10-425-114-11348	Sequence 11348, A
2	1166	49.0	1838	US-10-424-599-111422	Sequence 111422, A
3	833	35.0	1841	US-10-425-114-31883	Sequence 31883, A
4	653	27.5	2640	US-10-231-778-6	Sequence 6, Appl1
5	648.5	27.3	2700	US-10-437-963-91042	Sequence 91042, A

Result No.	Score	Query Match	Length DB	ID	Description
6	617.5	26.0	1004	US-10-437-963-11609	Sequence 11609, A
7	575.5	24.2	1070	US-10-424-599-141143	Sequence 141143, A
8	492	20.7	2439	US-10-425-114-30709	Sequence 30709, A
9	479	20.1	716	US-10-424-599-111421	Sequence 111421, A
10	472.5	19.9	443	US-09-837-604B-13507	Sequence 13507, A
11	472.5	19.9	443	US-09-837-604A-13507	Sequence 13507, A
12	435	18.3	558	US-10-389-048-21629	Sequence 21629, A
13	348.5	14.7	554	US-09-531-113-5697	Sequence 5697, A
14	348.5	14.7	554	US-09-531-113-5697	Sequence 5697, A
15	282	12.3	531	US-09-837-604B-940	Sequence 940, App
16	282	12.3	531	US-09-837-604A-940	Sequence 940, App
17	271	11.4	322	US-10-424-599-40645	Sequence 40645, A
18	264.5	11.1	6458	US-10-231-778-7	Sequence 7, Appl1
19	256.5	10.8	632	US-10-424-599-73456	Sequence 73456, A
20	255.5	10.7	453	US-09-837-604B-6414	Sequence 6414, App
21	255.5	10.7	453	US-09-837-604A-6414	Sequence 6414, App
22	235.5	10.7	453	US-10-437-963-97425	Sequence 97425, A
23	221	9.3	3235	US-10-170-235-25585	Sequence 25585, A
24	204.5	8.6	3900	US-10-144-771-9790	Sequence 9790, App
25	191.5	8.1	250	US-09-531-113-40606	Sequence 40606, A
26	191.5	8.1	250	US-09-531-113-40606	Sequence 40606, A
27	190	8.0	939	US-10-437-963-81637	Sequence 81637, A
28	184	7.7	274	US-10-389-048-16407	Sequence 16407, A
29	159	6.7	921	US-10-424-599-108137	Sequence 108137, A
30	153	6.4	3209	US-10-170-235-21113	Sequence 21113, A
31	140.5	5.9	350	US-10-424-599-9668	Sequence 9668, App
32	138	5.8	477	US-09-837-604B-52491	Sequence 52491, A
33	138	5.8	477	US-09-837-604A-52491	Sequence 52491, A
34	132	5.6	2250	US-10-369-493-45620	Sequence 45620, A
35	128	5.4	405	US-10-437-963-13671	Sequence 13671, A
36	124	5.2	3931	US-10-437-963-7175	Sequence 7175, App
37	123.5	5.2	564	US-10-437-963-98795	Sequence 98795, A
38	122	5.1	441	US-10-437-963-98795	Sequence 98795, A
39	119	5.0	3282	US-10-148-310-215	Sequence 215, App
40	119	5.0	3282	US-10-148-310-215	Sequence 215, App
41	118	4.9	782	US-10-437-963-33351	Sequence 33351, A
42	115.5	4.9	10640	US-10-304-095-5	Sequence 5, Appl1
43	115	4.8	2718	US-10-369-493-45577	Sequence 45577, A
44	114.5	4.8	2637	US-10-282-122A-16387	Sequence 16387, A
45	114	4.8	2892	US-10-282-122A-16387	Sequence 255, App

ALIGNMENTS

RESULT 1
US-10-425-114-11348
Sequence 11348, Application US/10425114
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E
APPLICANT: Tabaska, Jack E
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425, 114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 11348
LENGTH: 1722
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: 701055296_FLI
US-10-425-114-11348
Alignment Scores:
Pred. No.: 1,37e-116
Score: 1166.00
Percent Similarity: 65.65%
Best Local Similarity: 52.95%
Length: 1722
Matches: 242
Conservative: 58
Mismatch: 111

Query Match: 49.03% Indels: 46
 DB: 10 Gaps: 9
 US-09-890-220-2 (1-445) x US-10-425-114-11348 (1-1722)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 164 ATGTGCGGCAAAATTCCTCCGATACACATCCGGTGAAGAAAGAAATTCAGCTGATAG 223
 QY 21 AsnLeuLeuIleTyrcysLysProValArgLeuTyraAsnIlePheHisLeuArgSerLeu 40
 DB 224 AGCCTTTGATTATTCAGACCCGTGTGAACCTTTAAATAATTCCTGACCGCGCTCTT 283
 QY 41 GlysAnProSerPheLeuProArgCysLeuAsnTyrlsIleGlyAlaLysArgLysArg 60
 DB 284 CAAATCCCTCTTTCTTACGAGATGTTCCGTTATTAATAATGAAGCAAGCCGTAAAGG 343
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrlsAspCysAsnAsnThrLeuGln 80
 DB 344 AGCTTGAGAGCA--GGAAATGTGATTTCAATTATAGGACCCGTACAAACATCTCTTGA 400
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
 DB 401 AAGACTGAGTGAACCGAAGACTTTCTTCGCGTTTGGCTGTGATGACAGTGGAAACTTT 460
 QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
 DB 461 AAGGTTTGGGATTTCTTGTTCATCATCATGATCTTTCACTTGAAGTTCGGGTT 520
 QY 121 PheGluGluTyrglnThrValAsnValSerValLysLeuAsnSerPheIlePheGluGlu 140
 DB 521 ACTGAGATTAACCAAGAGTGAATGCTCTGTAATAATGAT-----ATATGAGATCA 574
 QY 141 GluLysSerAsp-----AspLysPheGluProPheSerLeuGlySerLysPro 157
 DB 575 GAGAAATGCTGTAAGAGTGAATTCACAAATCCCAACCTTTCTTCTGCTGTAAGACT 634
 QY 158 ArgLysArgArgGlnArgGlyLysArgAsnAsnThrArgArgLeuLysValCysPheLeu 177
 DB 635 CGAAACAGTGAAGAAAGAAAGCTGTTCAAATTTGAAAGCAACCAATGTAATAATCTCG 694
 QY 178 ProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAsp 197
 DB 695 GAGTTTGACTACCA-----GAAGCATATCAATGAATTTCTACAAAGAAAGATAT 745
 QY 197 ----- 197
 DB 746 GATATCTTATCTGCAAGAGAGAAATGTGTAGAACATCTGATGAGAGATTTTC 805
 QY 198 -----GlyAsnArgGlyLeuGlyTyrlsProGluAlaThrGluLeu 210
 DB 806 CCTAGTGAAGAAATGACGAGGAGAAATTTGGTCCGATCATCTGACCATGAGCAAC 865
 QY 211 AlaGlyLysPheGluMetThrSerAsnIleProPro-----AlaIleAlaHisSer 227
 DB 866 CTGAGCATGTGGAATTCAGTTCATTCACAGGTTTCAATTTGCCATGCCCAATCT 925
 QY 228 SerLeuAspAlaGly--AlaLysValIleLeuThrSerGluAlaValValProAlaThr 246
 DB 926 TCTGTGACCCCTGAATGTGTAATCAATATGTAAAGATGATCCGCTGCTGCTCT-- 982
 QY 247 LysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeuLeuLeuLys 266
 DB 983 AAAACAAAGAAAGCTAAGCATGATCATCATGATCAAGAACCGAATGCTGTGTGAGAG 1042
 QY 267 ArgGlnPheTyrlsSerHisArgValGlnProMetAlaLeuGluGlnValMetSerAsp 286
 DB 1043 AGACTTTCTTCTTACACACAGATGACCTATGCGCTAGAACCAAGTGTATTCACAC 1102
 QY 287 ArgAspSerGluAspGluValAspAspValAlaAspPheGluAspArgGlnMetLeu 306
 DB 1103 CGTCAATGTAAGAGCAAGTGTGATGACGATTCGAGATCTTGAAGATGAAGAGATGCTT 1162

QY 307 AspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeuThrAsnSerPhe 326
 DB 1163 GACGATTTTGTGATGATTTTCCAAAGATGAAAGAAACAGCTCATCTGCAACTCTTTT 1222
 QY 327 ValArgLysGlnArgValIleAlaAspGlyHisIleSerThrAlaCysGluAlaPheSer 346
 DB 1223 ATGAGAAAGCAAGGGTGTCTAGCATGATGATGTCGAGCCGTGTAGAGCATTTTCC 1282
 QY 347 ArgPheTyrgLysGluLeuHisArgTyrlsSerSerLeuPheThrPcysThrArgLeuPhe 366
 DB 1283 AAGCTTCATGAAAAGAGACTATGCTATCTCCAGCTTTATTTGGTGGAGGTTATTC 1342
 QY 367 LeuIleLysLeuThrPasnHisGlyLeuValAspSerAlaThrIleAsnAsnCysAsnThr 386
 DB 1343 ATGATCAAAACCTTGGAAATCAAGTCTTCTGATGCCGTACAAATGAACAACGTAGCAT 1402
 QY 387 IleLeuGluAsnCysArgAsnSerSerAspThrThrThrAsnAsnAsnSerVal 406
 DB 1403 GTATTAGATAGTTTACAGAAATGAGGATTCGGGTACAGAAATAAT----- 1447
 QY 407 AspArgProSerAspSerAsnThrAsnAsnAsnIleValAspHisPro 423
 DB 1448 TGAAGCCATTAAGAAAGAACTGATCAACACAGAAATATACATCTTCTCCG 1498

RESULT 2
 US-10-424-599-111422
 ; Sequence 111422, Application US/10424599
 ; GENERAL INFORMATION:
 ; APPLICANT: Ia Rosa, Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 111422
 ; LENGTH: 1838
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_71624C.1
 US-10-424-599-111422

Alignment Scores:
 Pred. No.: 1,52e-116 Length: 1838
 Score: 1166.00 Matches: 242
 Percent Similarity: 65.65% Conservative: 58
 Best Local Similarity: 52.95% Mismatches: 111
 Query Match: 49.03% Indels: 46
 DB: 10 Gaps: 9

US-09-890-220-2 (1-445) x US-10-424-599-111422 (1-1838)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluGluValIleSerThrAspGlu 20
 DB 258 ATGTGCGGCAAAATTCCTCCGATACACATCCGGTGAAGAAAGAAATTCAGCTGATAG 317
 QY 21 AsnLeuLeuIleTyrcysLysProValArgLeuTyraAsnIlePheHisLeuArgSerLeu 40
 DB 318 ACCCTTTGATTATTCAGACCCGTGTGAACCTTTAAATAATTCCTGACCGCGCTCTT 377
 QY 41 GlysAnProSerPheLeuProArgCysLeuAsnTyrlsIleGlyAlaLysArgLysArg 60
 DB 378 CAAATCCCTCTTTCTTACGAGATGTTCCGTTATTAATAATGAAGCAAGCCGTAAAGG 437
 QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrlsAspCysAsnAsnThrLeuGln 80
 DB 438 AGCTTGAGAGCA--GGAAATGTGATTTCAATTATAGGACCCGTACAAACATCTCTTGA 494
 QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100

[illegible]

```

Db      1542  TGAAGACCATAGAGAACTGACTGACACACAAAGAAATATATGACTCTTCGCG 1592

RESULT 3
US-10-425-114-31883
: Sequence 31883, Application US/10425114
: GENERAL INFORMATION:
: APPLICANT: Liu, Jingdong
: APPLICANT: Zhou, Yihua
: APPLICANT: Kovalic, David K.
: APPLICANT: Screen, Steven E.
: APPLICANT: Tabaska, Jack E
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: FILE REFERENCE: 38-21(5313)B
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 73128
: SEQ ID NO 31883
: LENGTH: 1841
: TYPE: DNA
: ORGANISM: Zea mays
: FEATURE:
: OTHER INFORMATION: Clone ID: UC-ZMLB73226G12.FL1
US-10-425-114-31883

Alignment Scores:
Pred. No.: 4.12e-80 Length: 1841
Score: 833.00 Matches: 180
Percent Similarity: 60.81% Conservative: 45
Best Local Similarity: 48.65% Mismatches: 83
Query Match: 35.03% Indels: 62
DB: Gaps: 8

US-09-890-220-2 (1-445) x US-10-425-114-31883 (1-1841)
QY 43 ProSerPheLeuProArgCysLeuAsnTyrIysIleGlyAlaLysArgIysArgIysSer 62
Db 459 CCMTCCTATTATCCAGAAATCATAGGTTGAGA----- 491
QY 63 ArgSerThrGlyMetValValPheAsnTyrIysAspCysAsnAsnThrLeuGlnIysThr 82
Db 492 -----GCTGGAAATGACTCTTTTATATACAGTACTACACAAATATACATGCAAAAGACT 545
QY 83 GluValArgIysAspCysSerCysProPheCysSerMetLeuCysGlySerPheIysGly 102
Db 546 GAAGTACTGAAAGATTTCCTGCGCCATTTCCTTGTCATGATGGAGAGCTTCAAGGCG 605
QY 103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGlnPheGlnPheLysLeuPheGlu 122
Db 606 CTAGAGAGCCCATTAATTCATCATCATGACCTGTTCCATAGAGTTGGATATCTTGAA 665
QY 123 GluIysGlnThrValAsnValSerValIysLeuAsnSerPhe-----IlePheGlnGlu 140
Db 666 GAGTGCACGCGCTGATATGTTAGTGTGAAGCGCTGATGCTGGAACAACTGATTTGTGGCA 725
QY 141 GluIysSerAspAspIysPheGlnProPheSerLeuCysSerIysProArgIysArg 160
Db 726 GAGGAGAGT---GATCCAAAGGCGATCAAAACATTTCTCTACTGCTCGAGGTTTAAAGAGCT 782
QY 161 ArgGlnArgGlyIysArgAsnAsnThrArgArgLeuLysValCysPheLeuProLeuAsp 180
Db 783 AGA-----AGCTTGGAAACCCACAGCTGAGAAA-----TTCAGGCATGTTCCAC 824
QY 181 SerProSerLeuThrAsnGlyThr----- 188
Db 825 CCACATATTATGAATCAGATCACATGACATGAGATGCCACAGGAGGCTGCAAGATGACTAT 884
QY 189 -----GluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlnIysPro 205
Db 885 GTGCAGAGGAAAGAAATGGGCTTTCTGTACCAAT----- 917
QY 206 GluAlaThrGluLeuAlaGlyGlnPheGlnLeuMetThrSerAsnIleProProAlaIleAla 225

```



```

Db 696 AGGATACAGATACATATATCATATACCTGGAAGTAACATTAAGCAATTCAGACAGCAAT 755
OY 61 -----
Db 756 ATCTTCCCTGATGTTCTGTTGCTAGACCTACTTCAAAATGCTTCTATAGAGACAT 815
OY 61 -----
Db 816 TCTCAATATATCGGTTGATGACGAGCCGTTGCTTACTTCTTATATGACTCTGGGAAT 875
OY 61 -----
Db 876 AATGACGCTGTGAAGCCACATTTGTTATCTGATCTGAGACTTAAATGCGACCCAA 935
OY 62 -----
Db 936 GCTTATGCTTACTTATCTTATCTTATGCGCGGTACCAAAAAAATAAGGCGCACT 995
OY 65 -----
Db 996 GGACAAAATCTTGTGAATAAGACTGTTCTGAGAAACATGTGACTACTCTCTCTCGCA 1055
OY 65 -----
Db 1056 AAGCTTGACGGAATGTTCTGCGGTAATAATCCGATCATTACTTAATTCATCTTTG 1115
OY 65 -----
Db 1116 GAGACTTGCGGATTAATTTGGGCGCATATAGTGAGACACCTATGATATGATATG 1175
OY 65 -----
Db 1176 AGCCAGGCTACTAGAGCCACACATTTCTTGACATGACATTTGCTTGCATTTTGTTC 1235
OY 65 -----
Db 1236 CGTAAGCTGATGCTATGCTTCCATATGCTGCAAGTAAGATATGACGACAGAGCT 1295
OY 65 -----
Db 1296 GGTGCAAAAGACATCTCAAAATCTCGTAAATTCCTTCTCATATAGTACGCTCCACCA 1355
OY 65 -----
Db 1356 TCCATTATGCTGCTATGTAAGGCGACGTGCTACTATGCTAATACATGAAGTCTTT 1415
OY 66 -----
Db 1416 TTGTACAGGCTAAGAGTGGAAATGCTCTTTACTATACAG-----AACACA-- 1463
OY 80 GlnysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySer 99
Db 1464 CAATAGAGTGAAGTAAGGAAATTTACTTCCCGCTTGGTGTGAGGCTGGGAAAC 1523
OY 100 PheLysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLys 119
Db 1524 -----TTCAAG 1529
OY 120 LeupheGluGluValArgGlnValAsnValSerValLysLeuAsnSerPheLeu 139
Db 1530 ATATCTGAAGACTACAGGCTGTTAATGTTAGCTGCTGAAGAAATTAACATGGAACAGAG 1589
OY 140 GluGluGlySerAsp---AspAspLysPheGluProPheSerLeuGlySerLysProArg 158
Db 1590 TTTGTGCGACGAGAGATTGATTAATAGCCATCGATCTTTACTACGATCAAGGTTTAA 1649
OY 159 LysArgArgGlnArgGlyGlyArgAsnAsnThrArgArgLeuValCysPheLeuPro 178
Db 1650 AAGAGTAAACAGAA-----ATACTTCCG 1673
OY 179 Leu-----AspSerProSerLeuThrAsnGlyThrGluAsnGlyLleThrLeu 195

```

```

Db 1674 GTTCCCGCTGACATGACACATTTATGGAATCAGATCA----- 1712
OY 196 AsnAspGlyAsnArgGlyLeuLysThrProGluValThrGluLeuValGlyGlnPheGlu 215
Db 1713 -----CCTAAATAACACAGAGGAGTGTGAGATGAT 1745
OY 216 MetThrSerAsnIleProPheAlaIleAlaHisSerSerLeuAspAlaGlyVal 235
Db 1746 GTCCAGAGAGAAATGAAACGCTTGATGATGATCTTAAATAATACATGCTAGCAAT 1805
OY 236 IleLeuThrSerGluValValProAlaThrLysThrArgLysLeuSerAlaGluArg 255
Db 1806 CATTCACAATCAGAA---TTCTGCAATTTGGGAAATCAAGGAACTATCAGCAATTCGA 1862
OY 256 SerGluValAlaArgSerHisLeuLeuGlnLysArgGlnPheThrHisSerHisArgVal 275
Db 1863 GCTGATCCAGAAATGCTACTCTGCAAAAAAGTCACTTATCATCTCTCATAGGCA 1922
OY 276 Gln----- 276
Db 1923 CAGACCCCTAGAGATGACAGAGTACCAACTTGTGTAAGGCGGATGCCATGTCGTA 1982
OY 277 PrometaLeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAsp 296
Db 1983 CCAATGACATTCGAAAGAGTTCTCTCAATATATATGTAAGTGAAGTATGATGAT 2042
OY 297 ValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlu 316
Db 2043 ATTGCTGATTTGGAGATGAGAGATGCTGATATTTTGTATGTTACAAAGATGAG 2102
OY 317 LysGlnPheMetHisLeuThrAsnSerPheValArgLysGlnArgValIleAlaAspGly 336
Db 2103 AAGCGATTATGACATGATGAGATTCATTTATGGAAGAAAGTAACTGACTACT 2162
OY 337 HisIleSerThrAlaCysGluAlaPheSerArgPheThrGluLysGluLeuHisArgTyr 356
Db 2163 CAGTACCTTGGGCTTGGAGCATTCCTCCGACATCATGAAACAACTTTAGAAAC 2222
OY 357 SerSerLeuPheThrCysTrpArgLeuPheLeuIleLysLeuTrpAsnHisGlyVal 376
Db 2223 TCCGCTTAACTATGGGAGTGGCGCATGTTATGATCAAACTCGGAATACAGTCTAC 2282
OY 377 AspSerAlaThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsn 393
Db 2283 TCTGCCCGCACATAGCACACCTCGCAAGAAATTCATGATACATAAAAAAT 2333

RESULT 6
US-10-437-963-11609/c
; Sequence 11609, Application US/10437963
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbezuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated wit
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437, 963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 11609
; LENGTH: 1004
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_17817C.1
US-10-437-963-11609
Alignment Scores:

```

Pred. No.: 6,136-57 Length: 1004
 Score: 617.50 Matches: 114
 Percent Similarity: 76.92% Conservative: 26
 Best Local Similarity: 62.64% Mismatches: 39
 Query Match: 25.97% Indels: 3
 DB: 10 Gaps: 1

US-09-890-220-2 (1-445) x US-10-437-963-11609 (1-1004)

```

QY 223 AAlaIleAlaHisSerSerLeuAspAlaGlyAlaVal-----IleLeuThrSer 239
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 844 TCTGAGACACAGCTCTGATGCTGTAATTCACAGGTACAGGATTCATTCACAGA 785
QY 240 GluAlaValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 784 CCAACAGTGTACATTTGGGAGACAGAAAGCTGTCTGTGAACAGCTATCCCA 725
QY 260 SerHisLeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnProMetAla 279
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 724 AATCGCAGCTCTACAAAGACGCGATCTTTCATTTCTCACAGGCTCAACCAATGGCA 665
QY 280 LeuGlnGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAsp 299
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 664 TTGGAGCAGATTCTTCAGATGATGATGAAGATGAAGTTGATGATGATGATGATGAT 605
QY 300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPhe 319
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 604 TTGAGAGATGAGAGATGCTGTATGATTTGTTGATGTTTCAAAAGACAGCAAACTTAAT 545
QY 320 MetHisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspLysHisLeuSer 339
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 544 ATGCAATGTGGAAATTCATTTGTTGGAACAAAGGTTACTAGCGATGGCATATTTCC 485
QY 340 TrpAlaCysGluAlaPheSerArgPheThrGlyLysGluLeuHisArgGlySerSerLeu 359
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 484 TGGGATGCGAAGCATCTCGCAAGTTTCATGAGCAAGAACTGTCAAAATCCAGCTCTCA 425
QY 360 PheTrpCysTrpArgLeuPheLeuLysLeuTrpAsnHisGlyLeuValAspSerAla 379
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 424 CTAATGCTGTGAGAGTTTATTATGTCAAACTCTGACACACAGCTACTGATGCGCA 365
QY 380 ThrIleAsnAsnCysAsnThrIleLeuGluAsnCysArgAsnSerSerAspThrThrThr 399
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 364 GCCAATGATGCTGCAACACAAATCTTTGAAAGCTACTGAAAGCAAGCTCGATCCAAAG 305
QY 400 ThrAsn 401
   |||
Db 304 AAAAAT 299

```

RESULT 7

US-10-424-599-141143/c ; Sequence 141143, Application US/10424599

GENERAL INFORMATION:

APPLICANT: La Rosa Thomas J

APPLICANT: Kovalic David K

APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53223)B

CURRENT FILING DATE: 2003-04-28

NUMBER OF SEQ ID NOS: 285684

SEQ ID NO 141143

LENGTH: 1070

TYPE: DNA

ORGANISM: glycine max

FEATURE: OTHER INFORMATION: Clone ID: PAT_MRT3847_98464C.1

US-10-424-599-141143

Alignment Scores: 2,666-52 Length: 1070

Score: 575.50 Matches: 123
 Percent Similarity: 63.14% Conservative: 38
 Best Local Similarity: 48.24% Mismatches: 63
 Query Match: 24.20% Indels: 31
 DB: 10 Gaps: 6

US-09-890-220-2 (1-445) x US-10-424-599-141143 (1-1070)

```

QY 145 AspAspLysPheGlnProPheSerLeuCysSerLysProArgLysArgGlnArgGly 164
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 1032 GATCCAAAGCTCAAAACATTTTCTTTGGAAGCCCTTAACCGTAAACGCAACAGCA 973
QY 165 GlyArgAsnAsnThrArgArgLeuValCysPheLeuProLeuAspSerProSerLeu 184
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 972 GACCAATCT-----TTGAAAATGCA-----GTGGCTTTAGAGCTTCCTTCT 928
QY 185 ThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAspGlyAsnArgGlyLeuGlyThr 204
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 927 GCAGAGGAGCTGAT-----ATTGAGAGAGAGATGATGAT----- 892
QY 205 ProGluAlaThrGluLeuAlaGlyGlnPheGluMetThrSerAsnIleProAlaIle 224
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 891 ----- 877
QY 225 AlaHisSerSerLeuAspAlaGlyAlaLysValIleLeuThr-----SerGluAlaVal 242
   |||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 876 ATTCGATCAGCTCCGATGAGACATCTGTTCAGTCAATGCTGACGTGATCAGCAAGTG 817
QY 243 ValProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArgSerHisLeu 262
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 816 CTTCAGTTTGGCAAGCAGAGAAAGTTGCAATGAGCGCTGACCCAGCAAAACAGTACC 757
QY 263 LeuLeuGlnLysArgGlnPheThrHisSerHisArgValGlnProMetAlaLeuGln 282
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 756 TTCTTGAGGAGCAGCAATTTTTCATTCACAAAGCTCAGCAATGCAATGCAATGCA 697
QY 283 ValMetSerAspArgAspSerGluAspGluValAspAspAspValAlaAspPheGluAsp 302
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 696 GTTCATATCGAATAAGATGCGAAGATGAAGTGAATGATGATGATGATGATGATGAT 637
QY 303 ArgGlnMetLeuAspAspPheValAspValAsnLysAspGluLysGlnPheMetHisLeu 322
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 636 CGAAGATGCTTGAAGAAATGTTGATGAGCAATGATGAGAGAGCACTTCATGATGAT 577
QY 323 TrpAsnSerPheValArgLysGlnArgValIleAlaAspGlnHisIleSerTrpAlaCys 342
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 576 TGGAACTCATTTGTTGGAGAGCATGTGTGATGTCAGATGTCATTTCAATGAGCATGT 517
QY 343 GluAlaPheSerArgPheThrGlyLysGluLeuHisArgLysSerSerLeuPheTrpCys 362
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 516 GAGGCTTCTCAAAATTCGATGACCTGAGTTTGTTCATCTCCCTCACTGGCAGAGGTGT 457
QY 363 TrpArgLeuPheLeuLysLeuTrpAsnHisGlyLeuValAspSerAlaThrIleAsn 382
   |||||.....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 456 TGGAGAAATGATTTATGTCAAATTAACAATCAATGATGCTTCATGATGCTGCGACATGAT 397
QY 383 AsnCysAsnThrIleLeuGluAsnCys---ArgAsnSerSerAsp 396
   .....|.....|.....|.....|.....|.....|.....|.....|.....|.....|
Db 396 GACTGTAAATTAATTTCTGAGCAATACCAAGGAGCAATTCAGAT 352

```

RESULT 8

US-10-425-114-30709 ; Sequence 30709, Application US/10425114

GENERAL INFORMATION:

APPLICANT: Liu, Jingdong

APPLICANT: Zhou, Yihua

APPLICANT: Kovalic, David K.

APPLICANT: Screen, Steven E

APPLICANT: Tabaska, Jack E

TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

FILE REFERENCE: 38-21(53313)B

CURRENT APPLICATION NUMBER: US/10/425,114
 CURRENT FILING DATE: 2003-04-28
 NUMBER OF SEQ ID NOS: 73128
 SEQ ID NO 30709
 LENGTH: 2439
 TYPE: DNA
 ORGANISM: Zea mays
 FEATURE:
 OTHER INFORMATION: clone ID: UC-2MFLB73067D07_FLI
 US-10-425-114-30709

Alignment Scores:
 Pred. No.: 1,286-42 Length: 2439
 Score: 492.00 Matches: 160
 Percent Similarity: 29.74% Conservative: 55
 Best Local Similarity: 22.13% Mismatches: 99
 Query Match: 20.69% Indels: 411
 DB: 10 Gaps: 12

US-09-890-220-2 (1-445) x US-10-425-114-30709 (1-2439)

OY 3 ArgGlnAsnCysArGAlaLysSerSerProGluGluValIleSerThrAspGluAsnLeu 22
 DB 283 CGTCAGCAGCTAAGAACTGATGCTCCAGATGAGAACTTACCCTGAAAGATTG 342
 OY 23 LeuIleTyrcysLysProValArgLeuTyraSniIlePheHisLeuArgSerLeuGlyAsn 42
 DB 343 GCTCGATTCGCAAGCCAGTCGAGCTCTACAAATATTATTCACGGCGAGCCATGAAAT 402
 OY 43 ProSerPheLeuProArGysLeuAsnTyrlsIleGlyAlaLysArgSerLeuGlyAsn 60
 DB 403 CCCCTTTTATCAAAATGCCCTCTTTAATATATCATGCGAGGAGGAAAGAGATT 462
 OY 60 ----- 60
 DB 463 CAGATAACCATATCATTCTTGAGAGTACAAATCTGAGTGCAGAAACATATATGCTTT 522
 OY 60 ----- 60
 DB 523 CCTCTTATGTTCTGTAGTAGACCCACTAGTACCTTTCAGTGAAGGCATTCTCCA 582
 OY 60 ----- 60
 DB 583 ATTATGCAATTCAGTCGGTGTGCTTCTTACTCTCTTATGGAACATGAAATAGGAC 642
 OY 61 ----- 65
 DB 643 AACAGTAGCTACATTCATCTCTGACGTGAGAGAGTTGTCTCAACCTCCCGTCTTGC 702
 OY 65 ----- 65
 DB 703 AACCATGATATATCTTTATTTAGCTGTGGCAGTTGACAAAGTAATGTGAGATATAC 762
 OY 65 ----- 65
 DB 763 TGTCTGGGAACCATGTGGAAGATCTCTCCAAATGCTTGAAGGAAATGCTCTGG 822
 OY 65 ----- 65
 DB 823 GGTAAATACCACTAATTTACTTGTCTCATCTTTGAGAGTTGTGTCAATTTAAGTTG 882
 OY 65 ----- 65
 DB 883 GGACATATTTGTGGAGTTGGCATTAAGTTACAAAGACCAAGCTCTTTAGAGCCAAA 942
 OY 65 ----- 65
 DB 943 TTTCTGAGCAAGACAGTGTCTGACATTTTGTCTCATTAAGCTTGATGCTGGGTTCA 1002
 OY 65 ----- 65
 DB 1003 TATAACTACACTATGATGTCGACAGAGGCTGTGTCAAGAGATATGTCTTGTCT 1062

OY 65 ----- 65
 DB 1063 CCATATAGTACTACTATTAATGANGTCCACCTTCGATTCATATCATATCATAG 1122
 OY 66 ----- 82
 DB 1123 TTAAGATCTGCAATGATCTTTTAATTTACAGTACTACATATATGCAAGAGACT 1182
 OY 83 GluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPheGly 102
 DB 1183 GAAGTCACGAGAGATTCTCTTGTCCATTTTGTCTATGTAGATGAGATGGAAGCTTCAAGGCT 1242
 OY 103 LeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeuPhe 121
 DB 1243 CTAGATGCAATTTAACTCATCATCATGATCTATTCACATATGAGTTGGGTGATATC 1302
 OY 122 -GluGluTyrglnThrValAsnValSerValLysLeuAsnSerPhe 136
 DB 1303 TGAAGATACAGGTTGTTAATGTTAGCTGAAGGCTGATGCTTGAGAAACAGAGTTC 1362
 OY 137 ----- 142
 DB 1363 TTCAAAGGTTGTTAAATGAGGCTCCACTTGTGGGCAAAACGGAAGCATTCGATG 1422
 OY 143 ----- 156
 DB 1423 TTCCTTAGTCCAGACGACCTCCGCTTTGTGATTCCTCTTCTTGTAG-TGCAG-AA 1480
 OY 156 sProArg 158
 DB 1481 GCCAAGATGAAGAAGCATTTGGCTATTCATATCCATCAAAATTCCTTACATAGTGA 1540
 OY 158 ----- 158
 DB 1541 TTGTGTACCTCAATGATTTTGTGTGATTAATGTTTACATGCTTAACATCATGCTTA 1600
 OY 158 ----- 158
 DB 1601 CTAATTTCTTAATGAGACACTCTTAAGTATTTGTAATGATGATCTAAGTGA 1660
 OY 158 ----- 158
 DB 1661 GCTTTTGGGAGGCGTTGATCCAGGCATCAAACTTTCTTATCGCTCAAGGTTAA 1720
 OY 159 ----- 177
 DB 1721 GAAGCTAGACGATCAAGAAACACATGAGAAATCAGGACATGACACTCATATATAT 1780
 OY 177 uProLeuAspSerProSerLeuThrAsnGlyThrGluAsnGlyIleThrLeuLeuAsnAs 197
 DB 1781 GGAATCAGGTTCCACTGAAAGATGAGGAGATCGAGAGCAACTTGTG-----CA 1831
 OY 197 pGluAsnArgGlyLeuGlyTyrrProGluAlaThrGluLeuAlaGlyGlnPheGluMet 217
 DB 1832 AGGGAGAGATGG----- 1844
 OY 217 rSerAsnIleProPheAlaIleHisSerSerLeuAspAlaGlyAlaLysVal----- 235
 DB 1845 ----- 1891
 OY 236 ----- 254
 DB 1892 CAGCAATCTTCCACCCACACAGACTACAGTTTGGAGACAAAGAAAGAACTTATC--GA 1948
 OY 254 uArgSerGluAlaArgSerHisLeuLeuGlnLysArgGlnPheTyrlHisSerHisAr 274
 DB 1949 GAGATCTACCCCTGAATCGCAACCTGCAAAAGAGAGTCTTCTCATATTCACAG 2008
 OY 274 gValGlnProMetAlaLeuGlnGluValMetSerAspArgAspSerGluAspGluValAs 294
 DB 2009 G----- 2009
 OY 294 pAspAspValAlaAspPheGluAspArgGlnMetLeuAspAspPheValAspValAsnly 314

```

Db      2009 ----- 2009
QY      314 saspsclulysglnphemethisleutrpasnserrphevalargylsgln-argvalillea 334
Db      2010 ----- GCGCAGAGCTGTAG 2023
QY      334 laspsclulysllesertrpalaCysgluAlaPheSerArpheyrglulysglnleu 354
Db      2024 CTGATGTCATACCTTGGCCCTGCGAGGCATTTCCCAAGTTGCATGAGCAGCAACTTA 2083
QY      354 lAArgTysSerleuPheTrpCysTrpArgLeuPheleuileuileuPheleuileu 374
Db      2084 TACAATAATCTGCTGCTGCGGGGTGGCGTTTCTCATGATTAACCTTGGAAACATA 2143
QY      374 lyleuValasPeralatrlleasnasCysAsnThrleuGlusnasCysArgAsn 394
Db      2144 ACATTTTATGATCCGCGCACTATGACATGCAATGCAATGCTTCAAAATTTTACAAGA 2203
QY      394 erSer 395
Db      2204 AAAGC 2208

```

```

RESULT 9
US-10-424-599-111421
; Sequence 111421, Application US/10424599
; GENERAL INFORMATION:

```

```

; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 111421
; LENGTH: 716
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_71623C.1
US-10-424-599-111421

```

```

Alignment Scores:
Pred. No.: 5,19e-42 Length: 716
Score: 479.00 Matches: 100
Percent Similarity: 72.46% Conservative: 21
Best Local Similarity: 59.88% Mismatches: 39
Query Match: 20.14% Indels: 7
DB: 10 Gaps: 3

```

```
US-09-890-220-2 (1-445) x US-10-424-599-111421 (1-716)
```

```

QY      1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu 20
Db      217 AGTGGCGGCAAAATTCCTCCGTCACACCATGCCGGGAGAAAGAAATTCACGCTGATG 276
QY      21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
Db      277 ACTCTTTTAAATTATGCAAGCGTGTGAACGTGACAAATATCTCTACCGCGCGCTT 336
QY      41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
Db      337 CAATAATCCCTTTCTTTAGAGATGTTGCTTATAAAATAAGACCAAGCGGTAAAG 396
QY      61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
Db      397 AGGTGAGAGCA---GGAATTGTGATTTCAATTATAGGATCACACACATCTTCCG 453
QY      81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuCysGlySerPhe 100

```

```

Db      454 AAAACTGAGTAGCAGCAAGACTTTCTGTCGCTTTGCTGATGCAAGTGTGAACTTT 513
QY      101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheGluPheLysLeu 120
Db      514 AAGGGTTTCGATTTATCTTCTGTCATACAGATCATTAACCTTGGGTTCTGGGTT 573
QY      121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGlu-Gl 140
Db      574 ACTGAAGATTACCAAGACAGTGAATGCTCCGCAAAATTAAT-----ATATTGCAATC 627
QY      140 uGluGlySerAspP-----AspLysPheGluProPheSerLeuCysSerLysPr 157
Db      628 AGAGATGTGCTGCGAGATGAGTATTCACAATCGCAAACTTCTTCTTGTTCAGAAC 687
QY      157 CArgLysArgArgGlnArg 163
Db      688 TCGAAACCGTAGAAGAAAG 706

```

```
RESULT 10
```

```
US-09-837-604B-13507
; Sequence 13507, Application US/09837604B
; GENERAL INFORMATION:

```

```

; APPLICANT: Byrum, Joseph R.
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Shukla, Hridayabhiraanjan
; APPLICANT: De la Pena, Robert C.
; APPLICANT: Bougri, Oleg
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604B
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 13507
; LENGTH: 443
; TYPE: DNA
; ORGANISM: Oryza sativa n1pnonbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3432-029-pl-K1-H2
US-09-837-604B-13507

```

```

Alignment Scores:
Pred. No.: 1,28e-41 Length: 443
Score: 472.50 Matches: 90
Percent Similarity: 80.60% Conservative: 18
Best Local Similarity: 67.16% Mismatches: 23
Query Match: 19.87% Indels: 3
DB: 5 Gaps: 1

```

```
US-09-890-220-2 (1-445) x US-09-837-604B-13507 (1-443)
```

```

QY      223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal-----IleLeuThrSer 239
Db      36 TCTGTGACACAGCTCTGTGATCTGTAATTCATTACAGGTGACAACTGTGTAGCA 95
QY      240 GluAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGluArgSerGluAlaArg 259
Db      96 CCACAGCTGTACAGTTGGGAAACACAAAGCTGCTGTGACAGACCTGATCCACA 155
QY      260 SerHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAla 279
Db      156 AATCGGCACCTCTTACAAAGCCAGTTCTTTCATCTCACAGGCGCTACACCAATGCA 215
QY      280 LeuGluGlnValMetSerAspArgAspSerGluAspGluValAspAspValAlaAsp 299
Db      216 TTGAGCAGAGTTTCTTCACATGCTGATGAGATGAGATGAGATGATGATGATGAT 275
QY      300 PheGluAspArgGlnMetLeuAspAspPheValAspValAsnLysAspLysGlnPhe 319
Db      276 TTTGAGATGAGAAATGCTTATGATTTTGTGATGTACAAAGACGAGAAACTTATT 335

```

QY 320 MethisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
 Db ATGCATATGTGTAATTCATTTGTCGGAACAAAGGCTACTACGCGATGCAATATTTCCC 395
 QY 340 TrpAlaCysGlnAlaPheSerArgPheTyrGluLysGlnLeu 353
 Db TGGCGATCGAAGCATTTCTCGAGTTTCATGACAGCAAGAACTT 437

RESULT 11

US-09-837-604A-13507
 ; Sequence 13507, Application US/09837604A
 ; GENERAL INFORMATION:
 ; APPLICANT: Byrum, Joseph R.
 ; APPLICANT: La Rosa, Thomas J.
 ; APPLICANT: Shukla, Hridayabhiranjan
 ; APPLICANT: De La Pena, Robert C.
 ; APPLICANT: Bougri, Oleg
 ; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
 ; FILE REFERENCE: 38-21(51892)B
 ; CURRENT FILING DATE: 2001-04-18
 ; PRIOR APPLICATION NUMBER: US 60/197,872
 ; PRIOR FILING DATE: 2000-04-19
 ; NUMBER OF SEQ ID NOS: 81288
 ; SEQ ID NO 13507
 ; LENGTH: 443
 ; TYPE: DNA
 ; ORGANISM: Oryza sativa nippondare
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: LIB3432-029-P1-K1-H2
 US-09-837-604A-13507

Alignment Scores:

Pred. No.:	1,28e-41	Length:	443
Score:	472.50	Matches:	90
Percent Similarity:	80.60%	Conservative:	18
Best Local Similarity:	67.16%	Mismatches:	23
Query Match:	19.87%	Indels:	3
Db:	7	Gaps:	1

US-09-890-220-2 (1-445) x US-09-837-604A-13507 (1-443)

QY 223 AlaIleAlaHisSerSerLeuAspAlaGlyAlaLysVal-----IleLeuThrSer 239
 Db TCTGAGACACGCGTCTGTGTAACCTGTAATTCATTACCGGTAACCAATCTTTCACCA 95
 QY 240 GlnAlaValAlaProAlaThrLysThrArgLysLeuSerAlaGlnArgSerGlnAlaArg 259
 Db CCAACAGTGTACAGTTGGGAAGACAAAGACCTGCTGTTGAACGAGCTGATCCCA 155
 QY 260 SerHisLeuLeuGlnLysArgGlnPheTyrHisSerHisArgValGlnProMetAla 279
 Db AATCGGACGCTCCACAAAACCCAGTCTTTCATCTCACAGGCGTCAACCAATGCA 215
 QY 280 LeuGlnGlnValMetSerAspArgAspSerGlnAspGlnValAspAspValAlaAsp 299
 Db TTGGAGCAAGTTTCTAGATCCTGATAGTAAGAGAGATGATGATGATCTCTCAT 275
 QY 300 PheGlnAspArgGlnMetLeuAspAspPheValAspValAsnLysAspGlnLysGlnPhe 319
 Db TTTGAGATGTAAGAAATGCTTGTGATTTGTTGATGCTTACAAAAGAGAGAAATTT 335
 QY 320 MethisLeuTrpAsnSerPheValArgLysGlnArgValIleAlaAspGlyHisIleSer 339
 Db ATGCATATGTGTAATTCATTTGTCGGAACAAAGGCTACTACGCGATGCAATATTTCCC 395
 QY 340 TrpAlaCysGlnAlaPheSerArgPheTyrGluLysGlnLeu 353
 Db TGGCGATCGAAGCATTTCTCGAGTTTCATGACAGCAAGAACTT 437

RESULT 12

US-10-389-048-21629

; Sequence 21629, Application US/10389048
 ; GENERAL INFORMATION:
 ; APPLICANT: Havukela, Ilkka
 ; APPLICANT: Sheenk, Michael A.
 ; TITLE OF INVENTION: Polynucleotides, Materials Incorporating
 ; FILE REFERENCE: 11000,1041U1C1
 ; CURRENT APPLICATION NUMBER: US/10/389,048
 ; CURRENT FILING DATE: 2003-03-13
 ; NUMBER OF SEQ ID NOS: 25129
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 21629
 ; LENGTH: 558
 ; TYPE: DNA
 ; ORGANISM: EucaIyptus grandis
 US-10-389-048-21629

Alignment Scores:

Pred. No.:	2,31e-37	Length:	558
Score:	435.00	Matches:	93
Percent Similarity:	68.29%	Conservative:	19
Best Local Similarity:	56.71%	Mismatches:	46
Query Match:	18.29%	Indels:	8
Db:	10	Gaps:	3

US-09-890-220-2 (1-445) x US-10-389-048-21629 (1-558)

QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGlnGlnValIleSerThrAspGlu 20
 Db ATGTGCAATCAAAAGTCTGTGACATTTTCTGATGATGACATGCACTGACGAGAGAG 133
 QY 21 AsnIleuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
 Db ATCTCTGCTGATATTCGACACCTGTGACGTATACAACTTTCACACCGCGCTGCATA 193
 QY 41 GlnAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
 Db CATATCTCTCTCTCTCTCGGAGATGTTGCTTCATAATACAAAGACGATAGCGCT 253
 QY 61 LysSerArg---SerThGlyMetValAlaPheAsnTyrLysAspCysAsnThrLeu 79
 Db AATAAGAGGTGAGGCGCTGAAATGATCTTTCATATAGGAGACCGCAACAGTGCAT 313
 QY 80 GlnLysThrGlnValArgLysAspCysSerCysProPheCysSerMetLeuCysGlySer 99
 Db CAAAAGGTGAAGTAACAGAAAGATTTCTTGCATCCCATTTGCTTAATGCAATGTGCAC 373
 QY 100 PheLysGlyLeuGlnPheHisLeuAsnSerHisAspLeuPheGlnPheGlnLys 119
 Db TTTAAGGCTCTAAGGCTTCATTTATGCTATCCGATGACTGTTCAACTTCGAGTTTGG 433
 QY 120 LeuPheGlnGlnTyrGlnThrValAsnValSerValLysLeuAsnSerPheIlePheGln 139
 Db GTTATGGAAGATATACAGCAGCAGTAATGTCTCTTAAATGTTAT-----ATGCT-GAG 486
 QY 140 GlnGlnGlySerAsp-----AspAspLysPheGlnProPheSerLeuCysSerLys 156
 Db ATCCGAGGTCTGACTGAAGAAGCAACCCATAATTTGACATTTCTTCTTGCCC-AAA 545
 QY 157 ProArgLysArg 160
 Db CCACGAGGCGG 557

RESULT 13

US-09-531-113-5697

; Sequence 5697, Application US/09531113
 ; GENERAL INFORMATION:
 ; APPLICANT: Byrum, Joseph R.
 ; APPLICANT: Heck, Gregory R.
 ; APPLICANT: La Rosa, Thomas J.
 ; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
 ; FILE REFERENCE: 38-21(15761)B

```

; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: jC-gmf102220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.: 6,64e-28 Length: 554
Score: 348.50 Matches: 76
Percent Similarity: 66.42% Conservative: 13
Best Local Similarity: 56.72% Mismatches: 26
Query Match: 14,668 Indels: 20
DB: 5 Gaps: 2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)
QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu 20
DB 193 ATGTGCGGCAAAATTCCTCCGGTACACCATGCGCGTGAAGAAATTCACGCTGATGAG 252
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 253 ACTCTTTAATTATTGCAAGCCTGTGTAACGTGTACAAATTTCTACCGCGCGCTCTT 312
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 313 CAAATCCCTCTTCTTCTTAGAGATGTTGCTTTATAAATAAGACCAAGCGCTAAAGG 372
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 373 AAGTTGAGAGCA---GGAATGTGATTTCAATATAGGATCCTACCAACATTCCTCGG 429
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 430 AAACGTGAAGTGACCGAAGACCTTTCTTCGCTTTGCTGATGAGGTGGCAGC--- 486
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLysLeu 120
DB 487 -----TTTAGCTTA 495

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
DB 496 CTGGAGAT--TACCAAGCAGTGAATGTCTCCGTGAATAATTAAT 536

RESULT 14
US-09-531-113-5697
; Sequence 5697, Application US/09531113
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: Heck, Gregory R.
; APPLICANT: La Rosa, Thomas J.
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(15761)B
; CURRENT APPLICATION NUMBER: US/09/531,113
; CURRENT FILING DATE: 2000-03-22
; NUMBER OF SEQ ID NOS: 48629
; SEQ ID NO 5697
; LENGTH: 554
; TYPE: DNA
; ORGANISM: Glycine max
; OTHER INFORMATION: Clone ID: jC-gmf102220073g03a1
US-09-531-113-5697

Alignment Scores:
Pred. No.: 6,64e-28 Length: 554
Score: 348.50 Matches: 76
Percent Similarity: 66.42% Conservative: 13

```

```

Best Local Similarity: 56.72% Mismatches: 26
Query Match: 14,668 Indels: 20
DB: 6 Gaps: 2

US-09-890-220-2 (1-445) x US-09-531-113-5697 (1-554)
QY 1 MetCysArgGlnAsnCysArgAlaLysSerSerProGluValIleSerThrAspGlu 20
DB 193 ATGTGCGGCAAAATTCCTCCGGTACACCATGCGCGTGAAGAAATTCACGCTGATGAG 252
QY 21 AsnLeuLeuIleTyrCysLysProValArgLeuTyrAsnIlePheHisLeuArgSerLeu 40
DB 253 ACTCTTTAATTATTGCAAGCCTGTGTAACGTGTACAAATTTCTACCGCGCGCTCTT 312
QY 41 GlyAsnProSerPheLeuProArgCysLeuAsnTyrLysIleGlyAlaLysArgLysArg 60
DB 313 CAAATCCCTCTTCTTCTTAGAGATGTTGCTTTATAAATAAGACCAAGCGCTAAAGG 372
QY 61 LysSerArgSerThrGlyMetValAlaPheAsnTyrLysAspCysAsnAsnThrLeuGln 80
DB 373 AAGTTGAGAGCA---GGAATGTGATTTCAATATAGGATCCTACCAACATTCCTCGG 429
QY 81 LysThrGluValArgGluAspCysSerCysProPheCysSerMetLeuGlySerPhe 100
DB 430 AAACGTGAAGTGACCGAAGACCTTTCTTCGCTTTGCTGATGAGGTGGCAGC--- 486
QY 101 LysGlyLeuGlnPheHisLeuAsnSerSerHisAspLeuPheGluPheLysLeu 120
DB 487 -----TTTAGCTTA 495

QY 121 PheGluGluTyrGlnThrValAsnValSerValLysLeuAsn 134
DB 496 CTGGAGAT--TACCAAGCAGTGAATGTCTCCGTGAATAATTAAT 536

RESULT 15
US-09-837-604B-940/C
; Sequence 940, Application US/09837604B
; GENERAL INFORMATION:
; APPLICANT: Byrum, Joseph R.
; APPLICANT: De La Pena, Robert C.
; APPLICANT: Shukla, Hridayabhiranjan
; APPLICANT: Bougri, Olegs
; TITLE OF INVENTION: Nucleic Acid Molecules And Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(51892)B
; CURRENT APPLICATION NUMBER: US/09/837,604B
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: US 60/197,872
; PRIOR FILING DATE: 2000-04-19
; NUMBER OF SEQ ID NOS: 81288
; SEQ ID NO 940
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Oryza sativa nipponbare
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3431-030-PI-NI-G11
US-09-837-604B-940

Alignment Scores:
Pred. No.: 9.44e-22 Length: 531
Score: 292.00 Matches: 49
Percent Similarity: 75.31% Conservative: 12
Best Local Similarity: 60.49% Mismatches: 20
Query Match: 12,288 Indels: 0
DB: 5 Gaps: 0

US-09-890-220-2 (1-445) x US-09-837-604B-940 (1-531)
QY 321 HisLeuTyrAsnSerPheValArgLysGlnArgValIleAlaAspLysHisIleSerTrp 340
DB 529 CATATGTGAATTCATTGTTCGGAACAAAGGCTACTAGCGGATGGCATATATCCCTGG 470

```

Fri Jun 20 09:03:41 2003

us-09-890-220-2.rtf

Page 12

QY	34	ALICSGIulnlaphesertraphetyrgIuysglIuenuHlstaGtYrserSerIuPhe	360
Db	469	GCATGGGAAGCATTCGCGAGTTTCATGGACAAGACCTTTACAAATTCACAGCTTACTA	410
QY	361	TRICSTYrPAtrGleuPheIuIeLysIuIeUrpAsnHlsgIuIeValaAspSerAlaThr	380
Db	409	TGtGTGTGGAGGTTTTTTATGTCACAACTCTGGAAACCACTACTACTGATGCGCAGCC	350
QY	381	ILAsnAsuScysAspThrIleuIuIuAsnGysaIyAsnSerSerAspThrIthrThr	400
Db	349	ATGAAATGCCCTGCACACACATTTCTTAAGGCTTACCTGAAGGAGAGCTCGATCCAAAGAA	250
QY	401	Asn 401	
Db	289	AAAT 287	

Search completed: June 20, 2003, 02:31:35
Job time : 982 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 13, 2003, 15:32:03 ; Search time 42 Seconds
(without alignments)
1018.567 Million cell updates/sec

Title: US-09-890-220-2

Perfect score: 2378

Sequence: 1 MCRONCRKSSPEEVISTDE.....INNKNYDNKDNRSRDYIK 445

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: pIR1:*
2: pIR2:*
3: pIR3:*
4: pIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	629	26.5	240	2 H71435	hypothetical prote
2	423	17.8	851	2 G71435	hypothetical prote
3	322.5	13.6	623	2 D71435	hypothetical prote
4	300	12.6	582	2 E84771	fertilization-inde
5	266	11.2	295	2 A71436	hypothetical prote
6	147.5	6.2	4550	2 T18440	hypothetical prote
7	140.5	5.9	1711	2 T18429	hypothetical prote
8	133.5	5.6	3848	2 T17414	TIPC protein - sli
9	132	5.6	749	2 A57276	MAD1 protein - yea
10	131.5	5.5	1245	2 D71613	GAF domain protein
11	130.5	5.5	699	2 T18426	hypothetical prote
12	128.5	5.4	1817	2 D71606	hypothetical prote
13	128.5	5.4	3844	2 T18402	asparagine/asparta
14	127.5	5.4	448	2 S05355	hypothetical prote
15	127.5	5.4	947	2 T08605	hypothetical prote
16	127.5	5.4	961	2 S67368	probable membrane
17	127.5	5.4	962	2 JCS808	G protein-coupled
18	126.5	5.3	1584	2 T18276	protein-tyrosine k
19	126.5	5.3	2150	2 S71629	sensory transducti
20	126	5.3	518	2 S23692	erythrocyte membra
21	125.5	5.3	1402	2 T17456	cell surface prote
22	124.5	5.2	686	2 A71607	Mtn3/PAG1P-like p
23	124	5.2	1619	2 T18499	hypothetical prote
24	123.5	5.2	1858	2 T18273	1-phosphatidylinos
25	123.5	5.2	2010	2 B71616	phosphatase (acid
26	123.5	5.2	3394	2 T18501	hypothetical prote
27	122.5	5.2	2269	2 T18472	hypothetical prote
28	122	5.1	720	2 T51007	hypothetical prote
29	122	5.1	1457	2 T14577	protein kinase yak

30	122	5.1	1714	2 E71609	Ser/Thr protein k1
31	120.5	5.1	608	2 T18437	hypothetical prote
32	120.5	5.1	1436	2 D71618	hypothetical prote
33	120	5.0	911	2 T18451	hypothetical prote
34	120	5.0	1119	2 T18491	hypothetical prote
35	120	5.0	1195	2 S38174	hypothetical purine nu
36	120	5.0	2380	2 E71604	hypothetical prote
37	119	5.0	1094	2 S46021	probable regulator
38	118.5	5.0	1278	2 A71609	probable secreted
39	118	5.0	3724	2 T18427	hypothetical prote
40	117.5	4.9	974	1 A40213	optic lobe develop
41	117.5	4.9	1075	2 S54067	probable membrane
42	117	4.9	590	2 S66956	hypothetical prote
43	117	4.9	2175	1 S03170	homeotic protein c
44	114.5	4.8	858	1 A42239	adenylate cyclase
45	114.5	4.8	1338	2 T18287	protein-tyrosine k

ALIGNMENTS

RESULT 1
H71435
Hypothetical protein - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 11-Jan-2002
C:Accession: H71435
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; D
.P.; Wedler, H.; Medler, E.; Wambutt, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; G
avanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgdomen
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.;
C.; Chalvatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis t
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: H71435
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-240 <BEV>
A:Cross-references: GB:Z97342; NID:G2245031; PID:G2245039
C:Genetics:
A:Map position: 4COP9-4G3845
Query Match 26.5%; Score 629; DB 2; Length 240;
Best Local Similarity 47.6%; Pred. No. 1.1e-41;
Matches 140; Conservative 2; Mismatches 0; Indels 152; Gaps 2;
OY 4 ONCRKSSPEEYISTDENILTYCKPVRLYNIFHLRSLGNPSFLPCLMYKIGAKRRK-- 61
|||||
DB 51 ONCRKSSPEEYISTDENILTYCKPVRLYNIFHLRSLGNPSFLPCLMYKIGAKRRKYYA 110
-----SRSTGMVFNKYDCNNLTOKTEVR 85
OY 62 -----
DB 111 FLFECSCSDMLFILLILWKLMLNYLSLSRYGHNMSRSTGMVFNKYDCNNLTOKTEVR 168
OY 86 EDCSCPFCSMLCGSKGLQFHLNSHDFFERFLFEYQYVNVSVKLSNLFEEEGSD 145
DB 169 -----
OY 146 DKFEPFLSKSPRRKRGGRNNTRRLKYCFLPLDPSFLTNGTENGITLLDNGNGLGYP 205
DB 169 -----GNRGGLGP 176
OY 206 EATEIAGOFEMTSNIPPAIAHSSLDAGAKVILTSBAVVPATKTRKLSARSEAR 259
DB 177 EATEIAGOFEMTSNIPPAIAHSSLDAGAKVILTSBAVVPATKTRKLSARSEAR 230
RESULT 2
G71435
Hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear-cress)
A:Variety: columbida
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998
C:Accession: G71435
R:Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirph
P.; Wedler, H.; Wedler, E.; Wandutt, R.; Weltzenegger, T.; Pohl, T.M.; Terry, N.; Gie
vanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Pulgomech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Reichman, S.; An
C.; Chaiwatzis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: D71435
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-851 <BEV>
A:Cross-references: GB:297342; NID:92245031; PID:6327517; PID:92245038
C:Genetics:
A:Map position: 4COP9-4G3845
A:Superfamily: Arabidopsis thaliana 70K hypothetical protein

Query Match 13.6%; Score 322.5; DB 2; Length 623;
Best Local Similarity 43.4%; Pred. NO. 3.1e-17;

	Matches	76;	Conservative	28;	Mismatches	38;	Indels	33;	Gaps	6;
Oy	237	LTSAAYVPAKTKRLKSERSEARSHLL-LQKRQFYHSHRYQPALQWMSDRSDEEV-	294							
Dd	6	LTTAKP-----AKRSKATSHYPLRKQRFYHSRTQPSLTDQWSDRSEVDVK	57							
Oy	295	-DDVADEDQMIDPFVDVK-DEKOMHHMNSPVKQVIADCHISMACEARSREYE	352							
Dd	58	NDDAAHLBESOMLGSDHEEIVAEKIKTAMSPVQAQRVLADAHLPACASRHLOB	117							
Oy	353	LHRSSLFWCKRFLIKLMHGLDVASTINCNTILENCNSSDTTTNNNSVD	407							
Dd	118	LRSNLSL-----DLGNTETTERAPV-----ATEATHREQDGD	151							
	RESULT 4									
E844771	Fertilization-independent seed 2 protein [Imported] - Arabidopsis thaliana									
C:Species:	Arabidopsis thaliana (mouse-ear cress)									
C:Date:	02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 02-Feb-2001									
C:Accession:	E84771									
R:LIn, X, J; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.										
M.; Koo, H.; Moffett, K.S.; Cronin, L.A.; Shen, M.; Varshen, S.E.; Umayam, L.; Tallon,										
euss, D.; Ntseman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter										
Nature 402, 761-768, 1999										
A>Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.										
A:Reference number: AB4420; MUID:20083487; PMID:10617197										
A:Accession: E84771										
A>Status: Preliminary										
A:Molecule type: DNA										
A:Residues: 1-582 >PRO>										
A:Cross-references: GB:AB002093; MID:g4263788; PIDN:AADI5448.1; GSPDB:GN00139										
C:Genetics:										
A:Gene: Atg35670										
A:Map position: 2										
Query Match	12.6%	Score 300;	DB 2;	Length 582;						
Best Local Similarity	19.6%;	Pred. No. 1.6e-15;								
Matches 117;	Conservative	54;	Mismatches 118;	Indels 308;	Gaps 13;					
Oy	13	EEVSTMDENILLYCKPYRLNYIFHLRSIGPSPSLPCIMNKIGAKRRKSRSGMVFNV	72							
Dd	46	DDVDVDVENIIRKIYPAYVKKLETRKN--	75							
Oy	73	KDCNNITLOKTEVRKEDSCFPCSMICGSFKLOPHLNSSHDLEFEFKLEEYO--TVNVS	130							
Dd	76	--	GLQLHLSSHDAKFEEFYRAEKDHGPVIDVS	106						
Oy	131	VKLNSFFEE--EGSSDDKEFPFSLCSKRP--KRORGGRRNTRLKYCELFDPSLTLNG	187							
Dd	107	VKSOTITFGVALKDVGGPQLSPILFCSKRNQRORDSNNAKLVALLMLDELDDLPRG	166							
Oy	188	TENGITILLDNCRNLGTPEA----	TELAGEEMTSNP--	PAIH	226					
Dd	167	TENDSTHVNDNWS--SPRAHSEKISDILTTOALAIAESSSEPKVPHVNDGVNSPPRAH	225							
Oy	227	SSLDAGAKVILITSEAVV----			243					
Dd	226	SSAEKNSTHVNDDDVVSPBRAHSLKNESTHVNDNISSPKAHSSKKNSTHNDED	285							
Oy	244	----PARKTRLSA-----ERSEARSHLLQKROF-----YHSRVOPMALE--	281							
Dd	286	VSPFPTRSSKETSIDLITTPQAIIVERSEPKEVVRARRKQLAKRYKARETOQAIAESSE	345							
Oy	282	----			281					
Dd	346	PKYLHAVNDENVSSPREAHSLEKASDILITTPQAIAESSEPKVPHVNDENVSSTPRAHSSK	405							
Oy	282	OWNSDRSDEEV-----			299					
Dd	406	KNSTKRVNDNVSPPTTRSSKTSDLITLTQPTIAESSSPKRVNHVNDNVSSTPRAHSS	465							
Oy	300	-----	FEDRQ	304						

Db 466 KKKSTRKNDNIPSPKTRSSKTSNLTATQPAKAPSEPKVTRRLKGRQFYHSQFW 525
 QY 305 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHSIMACEAFSEFEKELHRSLSFW 361
 Db 526 RLRLVGVSKKEKRYMYLWNLFFVRKQRYIADGHPWACEEFAKLHKEKKNSSPDW 582

RESULT 5

A71436

hypothetical protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

A:Variate: Columbia

C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 05-Dec-1998

A:Accession: A71436 B71436

R:Bean, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk

F.; Medler, H.; Medler, E.; Wandut, R.; Weitzenecker, T.; Pohl, T.M.; Terry, N.; Giel

avanagh, T.; Hempel, S.; Kotter, P.; Ertlan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.

Nature 391, 485-488, 1998

A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech

erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans

C.; Chalwatis, N.

A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal

A:Reference number: A71400; MUID:9812113; PMID:9461215

A:Accession: A71436

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-295 <BEV>

A:Cross-references: GB:297342; NID:92245031; PID:e327018; PID:92245041

A:Accession: B71436

A:Molecule type: DNA

A:Residues: 1-295 <BE2>

A:Cross-references: GB:297342; NID:92245031; PID:e327518; PID:92245040

C:Genetics:

A:Map position: 4COP9-4G3845

Query Match

Best Local Similarity 11.2%; Score 266; DB 2; Length 295;

Matches 70; Conservative 3; Mismatches 30; Indels 28; Gaps 4;

QY 278 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHSIMACEAFSEFEKELHRSLSFW 318

Db 1 MLDPEVDVNDKEQKFMHLMNSFVRKQRYIADGHSIMACEAFSEFEKELHRSLSFW 60

QY 319 FMHLMNSFVRKQRYIADGHSIMACEAFSEFEKELHRSLSFWCMLFLIKLHNGLVDS 378

Db 61 FMHLMNSFVRKQRYIADGHSIMACEAFSEFEKELHRSLSFWCMLFLIKLHNGLVDS 111

QY 379 ATINNCNTLE 389

Db 112 ALANVHTRE 122

RESULT 6

hypothetical protein C0425w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18440

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18440

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-4550 <LAW>

A:Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325396; PIDN:CAB11121.1

C:Genetics:

A:Map position: 3

A:Note: C0425w

Query Match

Best Local Similarity 6.2%; Score 147.5; DB 2; Length 4550;

Matches 19.8%; Pred. No. 0.018;

Matches 81; Conservative 64; Mismatches 150; Indels 115; Gaps 18;

QY 66 GMVFNVDKCN-----TLQTEVREDCSCFSCMGSLQFPLHSSHDLEFEER 119

Db 1820 GLTYTNLKNNNNNNNYMKTIKLT-----LCCT-----HSEHDTKAVMNR 1860

QY 120 L-----FEEXQTVNVS- -KLNSFIFEESGDDKFEPSLCSKPRKRGGRNTYRLK 173

Db 1861 LNTLNKKYKNNINCFYHYNSFLHSFPYSD-----EYIELIK-ERKEDIKEKTKRK 1915

QY 174 VCFPLDPSLTNGTEGTLTLDNGRGIGYEPATELAGQFEMTSNIPALHSSLDAGA 233

Db 1916 YCHOEDRDDNN 1967

QY 234 KYILSEAVVPAITKTRKLSAESEARSHTLLQKROFYHSHRQPMALDPEVSDSEDE 293

Db 1968 DQLINEED-----NKKRKINNLSHTHTNM-----SDHNKP-----KKKKNSEYQI 2011

QY 294 DDDV-----ADFEQMLDDEVD-----VNKDEKQPMHLMNSFVRKQRYIADG 337

Db 2012 NDDTKKKEHKKHEIDHKEDKQDNCNGIEKKDYKSKNNISKYVHFPERKNNI----- 2067

QY 338 ISWACEAFSEFEKELHRSLSFWCWRFLFLIKLHNGLVDSATINNCNTILENCRSSDT 397

Db 2068 -----KDEHK-----KEYNE-----RNDKMECDMLQNKRDNDHN 2099

QY 398 TTTNNNSVD-----RPSDNTNNNNIYDHPNDINNNNNVCKDNNSD 441

Db 2100 NNNNNNNDDNNND 2148

RESULT 7

T18429

hypothetical protein C0345w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000

A:Accession: T18429

R:Lawson, D.; Bowman, S.; Barrell, B.

submitted to the EMBL Data Library, August 1997

A:Reference number: Z18935

A:Accession: T18429

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-1711 <LAW>

A:Cross-references: EMBL:Z98547; NID:e1325376; PID:e1325381; PIDN:CAB11106.1

C:Genetics:

A:Note: PFC0345w

Query Match

Best Local Similarity 5.9%; Score 140.5; DB 2; Length 1711;

Matches 105; Conservative 95; Mismatches 204; Indels 165; Gaps 21;

QY 4 ONCRKSSPEEYISTD-----ENLLIYCKPVLYIIFHLRSLGNPSFLPCLN--YKIGA 56

Db 1054 OSNNNTSKDEHNISDSKSKEDTLMISRSKSYEYNNKLIQSTSNKS-----LNGAYENNL 1108

QY 57 KKKRSRSGVYVFNKDCNNLTQTEVREDCSCFSCMLCSFK-----GLQFHLNS- 110

Db 1109 FSGKKKKNGVLYLKDIEHNDIQDYPEDLNINCYNKYIENEKHLPLELEINLVSSD 1168

QY 111 -----HDFEEFEK--FEEXQTVNVS-----VKLNSEI 137

Db 1169 EKFGIKTKNDNNIITMKHQNHNLYDQKKHILFDTKNKSISQNNNNINSVITNHYE 1228

QY 138 FEEBSGDDKFEPSLCSKPRK-----RRORG-----RNNTR 170

Db 1229 VEKNNKDDRNNDYNT-CDKKKKITYIINISDKDIYHNNIITYKNEKEGIGNIHLNRNKD 1287

QY 171 RLKVCFLPDSF-----SLTNGTEGTLTLDNGRGIGYEPATELAGO-FEET 217

Db 1288 ITNFEILKLDGKKEFLDFKDSYIDCHNKEN-----ILNMTKNKEDHQIIVADKIFNET 1344

QY 218 SNIP-----PAIHSISLDAGAKVILTSEAVVPAITKTRKLSAESEA 258

```

Db      1345 NM1TM0NNK1YDDKKNHEKKCTHNDYIHNNMDI-----LSPSIRKN 1385
Oy      259 RSHLLLOKROFYHSHRVQPMALQ--VMSDRSEDEVDVADVEDROML-----DD 308
Db      1386 EENLFIDTYQ--KONRGIDYIMNRINILOEDDDDDNNNNNNNNNNNNKLLIFEXTKNDQ 1443
Oy      309 FVDVNDKQOFMHLMNSFVKORVADGHIWACAEAFKRYEKELHRYSSIFWCWRLFLI 368
Db      1444 MLHNNKNNLEGTSEEFSEFEKKKKI---KIKKNESYHKIDESILSNEKN-----NKVSL 1495
Oy      369 KLMNHGLVDSATINNCNTILENCRNSDPTTNNNNNSVDPDSNTNNNTYDHPNDIYN 428
Db      1496 LLINNNKSSVDNNKNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1555
Oy      429 K-----NNVDKNKDNNSRDVKIK 445
Db      1556 NNDSFSKDNMLINNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1584

```

RESULT 8

```

T1P3C protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T17414
R:Steger, J.T.; Iau, M.T.; Loomis, W.F.
submitted to the EMBL Data Library, July 1998
A:Description: Interaction of tip genes in early Dictyostelium discoideum development.
A:Reference number: Z18774
A:Accession: T17414
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 13848 <STE>
A:Cross-references: EMBL:AA079445; NID:93420744; PID:93420745; PIDN:AA031916.1
A:Experimental source: strain AX4
C:Genetics:
A:Gene: T1P3C
A:Introns: 72/3
C:Function:
A:Description: required for tip formation

```

```

Query Match      5.6%; Score 133.5; DB 2; Length 3848;
Best Local Similarity 19.7%; Pred. No. 0.18;
Matches 75; Conservative 47; Mismatches 103; Indels 155; Gaps 12;

Oy      76 NNTLQKTEVEREDSCPFCSMLCGSEFKLOPHLNSHDLFEFEFKLFEEYQTVNVSKINS 135
Db      2040 NNSNNOLPVEKEDSQ-----LQKYLISLEKGEF--LLNDHKISSPIKLLS 2083
Oy      136 FIFEEEGSDDDKFEPEFSLCSKPKRRQGRNNTRRLKVCFLPLDSPSLTNGTNGITLL 195
Db      2084 I-----GVDGLKSNIFSEF--PQK-----NOTALS 2105
Oy      196 NDGNRGLGY-----PEATELAGOFEKTSNIPPAIAHSSLDAGAKYILTSEAVVPATK 247
Db      2106 LDANNKAGYFNKNIGIWEPLIENWG--FSFTSN-----NSIEGGMWVNNESKIPLYINI 2157
Oy      248 TR-----KLSAERSEARSHLLQROFYHSHRVQPMALQVMSDSDSEVDVADVED 302
Db      2158 TKIFIDTSISTYQIMADYYSOQKD-----KKNKSNDYDNDDEITIEDY----- 2201
Oy      303 RQMLDPEVDVNDKQOFMHLMNSFVKORVADGHIWACAEAFKRYEKELHRYSSIFWC 362
Db      2202 EGVNNDLIKRPADQK----- 2216
Oy      363 WRFLIKLNMHGLVDSATINNCNTILENCRNSDPTTNNNNNSVDPDSNTNNNTYDHPNDIYN 422
Db      2217 -----NNNNNNKDNNSRDVKIK 445
Oy      423 PNDINKNNVNDKDNNSRDK 442
Db      2258 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2277

```

RESULT 9

```

A57276
MAD1 protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G3191; protein YGL086w
C:Species: Saccharomyces cerevisiae
C>Date: 08-Dec-1995 #sequence_revision 08-Dec-1995 #text_change 21-Jul-2000
C:Accession: A57276; S64093; S48246
J:Hardwick, K.G.; Murray, A.W.
J. Cell Biol. 131, 709-720, 1995
A:Title: Mad1p, a phosphoprotein component of the spindle assembly checkpoint in budd
A:Reference number: A57276; PMID:96042315; PMID:7593191
A:Accession: A57276
A:Molecule type: DNA
A:Residues: 1-749 <HAR>
A:Cross-references: GB:U14632; NID:9551091; PIDN:AAA91620.1; PID:9551092
R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.
submitted to the Protein Sequence Database, May 1996
A:Reference number: S64071
A:Accession: S64093
A:Molecule type: DNA
A:Residues: 1-749 <RIR>
A:Cross-references: EMBL:Z72608; NID:91322610; PID:e243297; PID:91322611; MIPS:YGL086
A:Experimental source: strain S288c
C:Genetics:
A:Gene: SCD:MAD1
A:Cross-references: SCD:S0003054; MIPS:YGL086w
A:Map position: 7L
C:Function:
A:Description: component of spindle assembly checkpoint which prevents cells from ini
C:Keywords: cell division control; nucleus; phosphoprotein

```

```

Query Match      5.6%; Score 132; DB 2; Length 749;
Best Local Similarity 21.3%; Pred. No. 0.029;
Matches 87; Conservative 57; Mismatches 143; Indels 122; Gaps 18;

Oy      89 SCFFCSMLCGS-----FKGIQPHLNSHDLFEFEFKLFEEYQTVNVSKINS 135
Db      40 SSFPLESPGSDVGTSGNSNRQIALQFLNTLQNEVIE--KLOLOKOTILKRYKA 98
Oy      136 FIFE--EEGSDDDKFE-----EPFSLCSKPKRRQGRNNTRRLKVCFLPLDSPSLTNGTE 189
Db      99 TIDELKALNDTKYLYESNDKLEQELSKLERSANSNMDK---CIEELRT----- 147
Oy      190 NGITLNDGNRGLGYPEATELAGOFE--MTSNIPPAIAHSSLDAGAKYILTSEAVVPATK 248
Db      148 ---TLQND-----LEWETLRQGYDLSKLTQVNOCDHPELEAS----- 183
Oy      249 RKLSEARSEARSHLLQ-----KROFYHSHRVQPMALQVMSDSDSEVDVADVED 302
Db      184 -----SHSLMKYERKIRQSVDIKIDQ---HQVM-----EKDELSSVKA 221
Oy      303 RQMLDPEVDVNDKQOFMHLMNSFVKORVADGHIWACAEAFKRYEKELHRYSSIFWC 352
Db      222 SKMINSHNYSYTER-----FNEITLKNKKIDQOYQYTKELALANQOANEL-----K 269
Oy      353 LHRYSIFWCWRLFLITLM-----HGL---VDSATINNCN-----TILENCRNSD 396
Db      270 LKQSDQSTFVKLENEKQNKLSQIAHLESQYENLOENLDLSKLTKEITYNDSDDDD 329
Oy      397 TTTTNNNNNSVDPDSNTNNNTYDHPNDIYNKNNVNDKDNNSRDVKIK 445
Db      330 NNVNNDNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 378

```

RESULT 10

```

D71613
GAR domain protein (cyclic nt signal transduct.) PFB0510w - malaria parasite (Plasmod
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: D71613
R:Gardner, M.J.; Teteltn, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.

```

Best Local Similarity 25.6%; Pred. No. 0.035;

RESULT 13

